

OY	501	CTGGGAGGTGGAGCGGTGGGCGGAGCGCTCTGGGGCTCAATTGATGTCTAGGGCATACGGA	560
Db	23766	CGCGGACATACCGGCGCTTCTCTCGTCCCGGTGGCGCTGCGACGTACCGACCGGGCGC	23825
OY	561	CGAGTGCAGCTTCTTGAGAGACGCGCTTTCGTGACAGCTGACGCGCGCGCTCAGCCA	620
Db	23826	ACTGGCGGCGGTGTGTA-----CGACGACGCGGCCACGCTGGCGTGGACGCGC	23874
OY	621	GGCCCGCTTACGGCGCTTGCCTGCGCGCCAGACGACAGACGGGCGCAATTGACACAA	680
Db	23875	CCGCGCTCTGGTGCCTCTGGGCGCGCGCTGGCGGAGACGCGACCGCG--CGAATGCGCGCGC	23933
OY	681	CGAGCTGAGAGAACAAACAGGTGGTGGCGAAGCTGGGGCGCATGGGCGCTGGGCGGCGCC	740
Db	23934	CTTGGCGGCCAGACACAGCGCGCGCGCGCACTGGTGTGACCTGGGCGCGCGCGCGCTT	23993
OY	741	GGGCGCACCGCGACGCGCACGCGCGCGCGCGCGCATTCGTGGTGGCGCGCGCGCGCGCC	800
Db	23994	CGAGCGCGCTGGTGTGCTTCTCTGGGTCTCGGAGTGGGGGGGGCGCGGCCAAGGGCGG	24053
OY	801	GCGGCGCGCGCTGCTCGCGCGCGCTACTGAGAGCGCGACCTGGCGTTTCATGCAACACGCG	860
Db	24054	CTACGCGGCGCGCACGCGCGCGCACTCGACGCGCTGGCGCGAACGCGCGCGCGCGCGGGGT	24113
OY	861	GGCGGAGTGAAGCGCTGGCGCGACCGCAAAAGTGGCTGGCGCACCGCGCGCGCGCGAGGCGG	920
Db	24114	GGCGCGCTTCTCGT---GGCTTGAAGCCCTCTGGGCGGAGGACGCGCGCGCGAGGTGC	24170
OY	921	CCGCAACGCTGTCTTCTTCGAGCGCGCGCTTGGCGGCGCGCGCGAGAGCGCTTCTGTGGAGGT	980
Db	24171	CGAGCGCGGAGTCTCTTACGCGCGCGCGCGCGGTGGCTCCCTGTGACCCGACCAAGCGGTGGG	24230
OY	981	GGGCGCTTCGAGGCGTGGCGCGCGCGCGCGCGCTGGCGCTTTCGCGATCACTGCGGACCC	1040
Db	24231	GACCTTCGCGCGCATGCTCTGAGAGCGCGCGCACGCGCTCGGTGCGGTTCGCGAGTGAAGTG	24290
OY	1041	GGGCGTGTCTACGCGCGCAACGAGCTGGCGCGCGCGCGACCGAGACGCGGTGCTGACCGCAAGA	1100
Db	24291	GAGCGGTTTGCGCGCTCTCTACACTGGAGTGGCGTCCCGCGGTACTTTGAGAGCAATCCC	24350
OY	1101	GTACTGGGTGGTGGCGCGC---GCCGGGCGCGTGGCGAGCGCGCGCGAGCGCGCTGAGCTT	1157
Db	24351	GGAAGTGAAGGGGTGGCGCGCGCGCGCGAATCTGCGCGCGAGACCGGAGATCTGACCACTTC	24410
OY	1158	CACGCTGCGCGCGCGCGCGCGACGTGC--TCTGGGCAATCAACGGGCGTTCGCGCGCGCGCC	1216
Db	24411	CGAATCTGTCGCGGAGCTGACCGGCGAGTCCGGCGCAAGAGCGCACGCACTCTGTGG	24470
OY	1217	TGCTGTGCTGTGACACCAACGACGCGCTCTGGGCG	1251
Db	24471	GCTGGTGGCGCACACGCGCGCGCGCGCTCTGCGAC	24505

RESULT 3
 US-08-804-198-1
 ; Sequence 1, Application US/08804198
 ; Patent No. 5945320
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgett, Stanley G.
 ; APPLICANT: Kuhnloss, Stuart A.
 ; APPLICANT: Rao, Nagarsaja R.
 ; APPLICANT: Richardson, Mark A.
 ; APPLICANT: Roseck, Paul R., Jr.
 ; TITLE OF INVENTION: PLATENOULIDE SYNTHASE GENE
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSSEE: PAUL R. CANTRELL 1138
 ; STREET: LILLY CORPORATE CENTER
 ; CITY: INDIANAPOLIS
 ; STATE: IN
 ; COUNTRY: USA
 ; ZIP: 46285

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CANTRELL, PAUL R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 44377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 350..14002
FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036
FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284
FEATURE:
NAME/KEY: CDS
LOCATION: 31329..36071
FEATURE:
NAME/KEY: CDS
LOCATION: 36155..41830
US-08-804-198-1

Query Match 4.8%; Score 79.8; DB 2; Length 44377;
Best Local Similarity 4.5%; Pred. No. 3.2e-06;
Matches 549; Conservative 0; Mismatches 667; Indels 19; Gaps 5;

21 GACCTTCGACAGCCCGAAGCCACCCGCGCGCTCTCTGAGCACCCGCGCTGCGCGCC 80
23286 GCCGTCGCGCTGACCCGACCCCTCGCCCTCTGTCAGAGGCTTGACCTCGGACGAC 23345
81 CGGCGCCGACGACGCGCGGCTCTGCGGCGAGGCGCGGCTTCCAGCGCGACGACAAAG 140
23346 CGACCGCTGTGACCTTCACCTGCGCGCGGCTGCGCACCGCCCCGACGAACTGCGTG 23405
141 CAAGAAGTGCAGCTGACGCGACGCGACCTGCGCGCGGACACAGCGCGCAAGCTTTCGAA 200
23406 CACGCGCGGTGCCAGCTGTGTGGGGCTGTGGGCGGGGTGGCGCGCTGGAAGTCCCGAAGT 23465
201 TGGCGTACGTTACGACGAGCGGCAATCGGCTGTACAGACAGTGGCGCTGCTGT 260
23466 GTGGGCGGCGCTCATGACCTTCCCGCGCGGCGCGACGCGCGGGTCTGGAACGTTCGCG 23525
261 GCGGTGCGGCTGTGAGCTGAGAGCGGCGCGCTGCGCTTCCGCTTCCAGCGCGCAAGTCCGTC 320
23526 CGGGTCTTGCAGCAACCGCGCGGCGAGACAGAACTGCGGTACGAGTGGCGGCGCTTT 23585
321 GCTCATGAGCGCCAGGACATCCCAAGTACGCTGCGCGGACCTGTGACGCGGCGCGG 380
23586 CGGCGCGGGGTCTTGGGGAACCCGCGCACTCCCGGCCCCCGGCGCGGCGCGCGG 23645
381 CTACTGGGCGCAAGGCACTGCGCGGAGAACTGGCGCTGCGGACACGATGCTGGCTTACTG 440
23646 CACGTCCTCATGCGCGGCGACCTTCAAGAGGTGGCCGCGCGCACTGATCGGTCCCTCT 23705
441 GCGCGACCGCACGCGCGCTGTCTTACAGCTGAAAGACAGCGGACCGGTCTCTTCCA 500

```

```

Db      23706  CGAGAGCGCGCGGACCGCGTGTGTGCGCGGAGCCGACGCCCCCGCACAGGCCCGC 23765
QY      501    CTGCGGCGGTGGCGGGGGGCGCGGCTTGGGGGCTCATTTGATGTCTAGGACATCAACGA 560
Db      23766  CGCGGAGCTGACCGGGGCTCTCCCTCGTCCCGTGGCGGTGAGAGTCAACGACCGCGCGC 23825
QY      561    CGAGGTGACAGCTTCTGAGAGAGCGCTTGTGTGACAGCTGACCGCGCGCGCTCTAGCGCA 620
Db      23826  ACTGGCGGCGCTGTGCA-----CGAGCAGCGCGCGCACCGTGGCGGTGACAGCGC 23874
QY      621    GGGCGGCTTCAAGGCGCTTGTCTGCGCGCCAGACGACAGAGCGCGCAATTGACAAACA 680
Db      23875  CGCGCCCTGTGTGCGCGTGGCGCGCTGCGGAGACGCGACCCCGG-CGACATCGCGCGCG 23933
QY      681    CGAGCTCGAGAACAAACAGGTGTGAGCAAGCTGGGCACTGGGCGCTGGAGCGCGCGCC 740
Db      23934  CTTGCGCGCGCAACACAGCGCGCGCGCGCGACCTGTGTGACCTGCGCGCGCGCGCGCG 23993
QY      741    GGGCCGACCGCGCACGCGACGCGCGCGCGCGCGCATTTCCGTGCGGCGCGCGTGAAGCGCC 800
Db      23994  CGACGCGCTGTGTGTCTGTCTCTCGGTCTCGGAGTGTGGGCGCGCGCGCGCGCGCGCG 24053
QY      801    GCGGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 860
Db      24054  CTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 24113
QY      861    GCGCGGAGTGAAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 920
Db      24114  GCGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24170
QY      921    CGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 980
Db      24171  CGAGGCGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24230
QY      981    GGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1040
Db      24231  GACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 24290
QY      1041  GGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1100
Db      24291  GAGCGGTTGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24350
QY      1101  GTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1157
Db      24351  GAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 24410
QY      1158  CAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1216
Db      24411  CGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 24470
QY      1217  TGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1251
Db      24471  GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 24505

```

```

RESULT 4
US-09-105-537-34
; Sequence 34, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 4689
; TYPE: DNA

```

```

; ORGANISM: Streptomyces venezuelae
US-09-105-537-34

```

```

Query Match      4.7%; Score 79.4; DB 3; Length 4689;
Best Local Similarity 44.1%; Pred. No. 3,5e-06;
Matches 427; Conservative 0; Mismatches 536; Indels 6; Gaps 2;

```

```

QY      317  CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 376
Db      2723  CGCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2782
QY      377  CGGCTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 436
Db      2783  CGGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2842
QY      437  ACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 496
Db      2843  AGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2902
QY      497  TCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 556
Db      2903  ACAGAGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2962
QY      557  CGGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 616
Db      2963  CGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3022
QY      617  GCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 676
Db      3023  GCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3082
QY      677  ACAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 736
Db      3083  GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3142
QY      737  CCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 795
Db      3143  CCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3202
QY      796  CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 855
Db      3203  CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3262
QY      856  CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 910
Db      3263  TCCGACGACCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3322
QY      911  CGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 970
Db      3323  CCAACCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3382
QY      971  TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1030
Db      3383  GCAAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3442
QY      1031  CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1090
Db      3443  GGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3502
QY      1091  ACCGCAAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1150
Db      3503  CCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3562
QY      1151  TCAAGTTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1210
Db      3563  AGCGCATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3622
QY      1211  GCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1270
Db      3623  GCGATCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3682
QY      1271  GCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1279

```

Db 3683 GCGGCGCCG 3691

RESULT 5

US-09-105-537-5

Sequence 5, Application US/09105537A

Patent No. 6265202

GENERAL INFORMATION:

APPLICANT: Sherman, D.H.

APPLICANT: Liu, H.

APPLICANT: Xue, Y.

APPLICANT: Zhao, L.

TITLE OF INVENTION: DNA encoding methymycin and pikromycin

FILE REFERENCE: 600,438U1

CURRENT APPLICATION NUMBER: US/09/105,537A

CURRENT FILING DATE: 1998-06-26

NUMBER OF SEQ ID NOS: 43

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 5

LENGTH: 36778

TYPE: DNA

ORGANISM: Streptomyces venezuelae

US-09-105-537-5

Query Match 4.7%; Score 79.4; DB 3; Length 36778;
Best Local Similarity 44.1%; Pred. No. 3.7e-06;
Matches 427; Conservative 0; Mismatches 536; Indels 6; Gaps 2;

```
QY 317 CGTCGCTATGAGCCGCCAGGACATCCCAAGTACGCTGCCGCGGACCTGTCTACGCGGC 376
DB 29713 CGCCACACAGCGCCGCCAGCACTGACGCTTACCGGCTGAGTGAAGCCGCTGACGCGCT 29772
QY 377 CGGCTACTAGGCGCAAGGACCTGCGCGGAGAACTGCGCGTGGCCACACAGGTGCTGCGCT 436
DB 29773 CGGCGCAAGCGGACCTGTCTCGGCGCGGTGATGTGCGCTCGGAGCGGACCAAGAACCG 29832
QY 437 ACTGGGCGGACCGGACCGCGCGGTGTTCTACAGGTGAACGACGCGGACGCGGTGCTT 496
DB 29833 AGCTCTGCGGCGCGCTGAAGCGCGGAGCGGAGCGGAGTCACTACTGAAAGCGGCGCG 29892
QY 487 TCCTATGCGGCGGTGCGCGGTGCGCGCGCGCGCTCTGCGCGCTACTTATGTCTACGCACTA 556
DB 29893 ACGAGACCGGTGAGGCGCTGCGCGCGCGCTGACCGCACTGACGACCGCGACGCGCTTCA 29952
QY 557 CGGACGAGGTGACGCTTCTGAGAGGCGCTTCTGCTGACACGCTGACGCGCGCGCGCTCA 616
DB 29953 CGGCGGTGTCTGCTCTCTGACGACCTGTGCTGACAGGTGCGCTGCGGTGACGCGCACTG 30012
QY 617 GCGAGCGCGCTTCAAGCGCTGCGCTGCGCGCGCGGACGACGACGCGCGCACTTCAACA 676
DB 30013 GCGAGCGCGGATCAAGGCGCGCGCTGTGCTGCTGACCGACGCGCGCGGTCTGCTGCGAC 30072
QY 677 ACAACGAGCTGAGAAACAACAGGTGAGGCGCAAGCTGCGCGCACTGCGCGCTGCGCGCG 736
DB 30073 GTCTGACACCGCGCGCGCGCGCGCGCGCGCGCATGCTCTGCGCGCGCGCGCTGCTG 30132
QY 737 CCGCGCGCGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCATTCGTCGCGCGCGCGCGG-AG 795
DB 30133 CCGCTTGAAGACCGCGAGCGCTGCGCGCGCGCGCGCTGTGACCTTCCCGCGCGCGCGATGCC 30192
QY 796 CGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 855
DB 30193 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 30252
QY 856 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 910
DB 30253 TCCTGACACCGCGAGTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30312
QY 911 CGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 970
DB 30313 CGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30372
QY 971 TCGTGAAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1030
```

```
DB 30373 GCGCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30432
QY 1031 CGTCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1090
DB 30433 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30492
QY 1091 ACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150
DB 30493 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30552
QY 1151 TCAGCTTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1210
DB 30553 ACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30612
QY 1211 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1270
DB 30613 GCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30672
QY 1271 GCGTGGCGG 1279
DB 30673 GCGGCGCGG 30681
```

RESULT 6

US-09-320-878-19

Sequence 19, Application US/09320878A

Patent No. 6117659

GENERAL INFORMATION:

APPLICANT: ASHLEY, Gary

APPLICANT: BETLACH, Melanie C.

APPLICANT: BETLACH, Mary C.

APPLICANT: MEDANIEL, Robert

APPLICANT: TANG, Li

TITLE OF INVENTION: RECOMBINANT NARHONOLIDE POLYKETIDE SYNTHASE

FILE REFERENCE: 30062202120

CURRENT APPLICATION NUMBER: US/09/320,878A

CURRENT FILING DATE: 1999-05-27

EARLIER APPLICATION NUMBER: CIP OF 09/141,908

EARLIER FILING DATE: 1998-08-28

EARLIER APPLICATION NUMBER: CIP OF 09/073,538

EARLIER FILING DATE: 1998-05-06

EARLIER APPLICATION NUMBER: CIP OF 08/846,247

EARLIER FILING DATE: 1997-04-30

EARLIER APPLICATION NUMBER: 60/119,139

EARLIER FILING DATE: 1999-02-08

EARLIER APPLICATION NUMBER: 60/100,880

EARLIER FILING DATE: 1998-09-22

EARLIER APPLICATION NUMBER: 60/087,080

EARLIER FILING DATE: 1998-05-28

NUMBER OF SEQ ID NOS: 34

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 19

LENGTH: 38506

TYPE: DNA

ORGANISM: Streptomyces venezuelae

US-09-320-878-19

Query Match 4.7%; Score 79.4; DB 3; Length 38506;
Best Local Similarity 44.1%; Pred. No. 3.7e-06;
Matches 427; Conservative 0; Mismatches 536; Indels 6; Gaps 2;

```
QY 317 CGTCGCTATGAGCCGCCAGGACATCCCAAGTACGCTGCCGCGGACCTGTCTACGCGGC 376
DB 27855 CGCCACACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 27914
QY 377 CGGCTACTAGGCGCAAGGACCTGCGCGGAGAACTGCGCGCTGCGCGGACACAGGTGCTGCGCT 436
DB 27915 CGGCGCAAGCGGACCTGTCTCGGCGCGGTGATGTGCGCGCGGAGCGGACGCGGCGCG 27974
QY 437 ACTGGGCGGACCGCGCGCGCGCGCGGTGTTCTACAGGTGAACGACGCGGAGCGCGGTGCTT 496
DB 27975 AGCTCTGCGGCGCGCTGAAGCGCGCGGAGCGGAGGTGACGTAAGCGCGGCGCGG 28034
```


Accession	Sequence	Position
Dd	28635 CCGCGTCACCATGCGCGCTGCGACGTCCGCCAACCCCAAGGCATGCGCACCTCTCTCG	28694
Qy	1151 TCAGCTTCACGCTGCGGCGCGGCGGCGACGCTGCTCTGCGCATCAACGGGCGGTGCGCGG	1210
Dd	28695 ACGGCATCCCGCGCGGAGACGCCCTCTACCGCGCTGTGTCCACACCGCGCGGACCGCGGCG	28754
Qy	1211 GCCGCTGCTGTGCGGTGACACCAACGACGAGCGGCTCTGAGGCTTCTTTCGCGGTGCGGCGG	1270
Dd	28755 GCGATCCGCGTGGACGTCAACCGGCGCGGAGGACATCGCCGCACTCTGGGCGGAGACGA	28814
Qy	1271 GCGTTCGCGG 1279	
Dd	28815 GCGGCGCG 28823	

```

RESULT 8
US-09-657-440-19
; Sequence 19, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062200210
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-657-440-19

```

Query Match	Similarity	4.7% 44.1%	Score 79.4	DB 4	Length 38506
Best Local Match	Similarity	44.1%	Pred. 0.37e-06		
Matches 427	Conservative	0	Mismatches 536	Indels 6	Gaps 2
Db	317	CGTGCCTCATGAGCGCCCAAGACATCTCCCAAGTACGCGCTTCCCGGACCTTGATCAACGCGAC	376		
Db	27855	CGCCACACAGCGCGCCGACGACTGAGGCTTACCGCGCTGAGTGAAGCCGCTACGCGCT	27914		
Qy	377	CGGGCTACTGGGSCCAAGGCACTGCGCCGAGAACCTGGGCGCTGCGGACACGAGTCTGGCT	436		
Db	27915	CCGGCCAGCGCGGACCTTCTCGGGGGTGGATCTGTCGCGCTGGAGAGGACCAAGACCG	27974		
Qy	437	ACTGGGCGGACCGGACCGGCGCGCTGTTTCAACAGCTGAACGACCGGCGAGCGAGTCTCT	496		
Db	27975	AGCTGCTGAGGCGGCGCTGAAGCGCGCGGAGGAGGAGTGCACGTACGTGAAGCGGGGCGG	28034		
Qy	497	TCCACTGGCGGCTGGCCGCTGGGCGGCGCGCTCTTGGCGCTCATATTGATCTACGCACTCA	556		
Db	28035	ACGACGACCGTGAAGGCGCTTCCGCGCGCGGCTCAACCGACATGAAGACCGGGGCGCTTCA	28094		
Qy	557	CCGACGAGAGTTCACGCTTCTTGGAGAGCGCTTCCGCTGACACGCTGACGCGCCGCGGCTCA	616		
Db	28095	CCGGCGTGGTCTCGCTCTCTGACGACCTTCGTGCACAGTGGCGCTGGTGAAGGCACTCG	28154		
Qy	617	GCGAGGCGCGCTTCAGCGCGCTGCTCGCGCCGACGACGACGACGCGGCGCACTTGACAC	676		
Db	28155	GCGACGCGCGAATCAAGGCGCGCGCTGTGGTCCGCAACCGAGGCGCGAGTCTCGTGGAC	28214		
Qy	677	ACAAAGAGCTGAGAAACAACAGTGGTGGCGMAAGCTGGGCGACCTTGGCGCTGGGCGCGG	736		
Db	28215	GCTTCGACACCCCGCGGACCCCGACCGGGCGCAAGCTCTGAGGCGCTTGGCGGCTGGTG	28274		

OY	737	CCCCGGGGCCCAACGGCCAGCGGAAACGGCCGGGCGCGCGCCATTCCTGTGCGGGGCCCCGGG-AG	795
Db	28275	CCCTTAGGCAACCCCGAAACGGCTGGGGCGGCGCTGTCACTCCCGGCCACGCCCGAATGCCG	28334
OY	796	CGCCCGCGCGCCCGCGTGTTCGCGCGGCGCTAATTGAGAGCGAACCTTCGCTTTCACACGCA	855
Db	28335	CCGCGCTTCGGCCCACTCTGTCAACCGCACTCTCCCGGCGCCACCGCGAGAGCACAGATGCGCA	28394
OY	856	CGCGGGGCCCCGACGTGAAGCTTGTGGC-----GACCGCGAAATGGCGCTTCGCAACCGCGC	910
Db	28395	TCCGCACACACCGGACCTTCACACGCCCGCGCGCTCGCCCGCACCCCTTCACCGACGTTCGCG	28454
OY	911	CCGACCGGCGGCGCGCAGCTGTGTCTTCGCGAGCGCCCGTGGCGCGCGCGCGGAGCGCTCT	970
Db	28455	CCACCCCGGACTGGACAGCCCCCAACGGACACGCTCTTATCACCGCGGCGCACCGAGCCCTTCG	28514
OY	971	TCTGTGAGAGTGGGCGCGTCCGCGGGCTGTGAGCGCGCCCGCGCGCTGGCGCTTTCGGCATCACTG	1030
Db	28515	GCAAGCCACGGCGCGCAGCTGTGATGGCCCAACAAGGACCGAACAACCTCTCTGTATAGCG	28574
OY	1031	CGTGCAGACCCGGGCGGTACTACGGCCCCAAGAGCTGCGCGCGCACCAAGACGGCGTCTCG	1090
Db	28575	GCAAGCGCGGAAACAAGCCCCCGGAGGCGAACCACTCACCCCGCGAATCAACCGCATCCGGCG	28634
OY	1091	AACCGCAAAAGATACTGTGGTGTGTGTGGCGCGCGCGCGCGCGTGGCGAGACCGCGCGAACCGCC	1150
Db	28635	CCCGCGTCAACCATTCGCGCGCTGTGGAGCGTGCACCCCAACGCGACATGCGCAACCTTCTCG	28694
OY	1151	TCAAGCTTCACGCTGTGGCGCCCGGCGGCGAACGTGTCTCTGTGGACATCAACGGGCGTCCGCGCG	1210
Db	28695	ACGCGCATCCCCCGCGAGAACGCCCTCTACACCGCGCGTGTCCACACACCGCGCGCGCACCGGGG	28754
OY	1211	GCGCGCTGTGTGTGTGACACCAACGCAAGCGCTTGTGGGCGTTCTTTCGCGCGTGGCGGCGCG	1270
Db	28755	GCGATCCGCTGTGACGTCAACCGGCGCGGAGGACATCCGCGCATCTCTGGCGCGGAGACGA	28814
OY	1271	GCGTCCGCGG-1279	
Db	28815	GCGGCGCGCG-28823	

```

1  RESULT 9
2  US-08-403-852D-6
3  / Sequence 6, Application US/08403852D
4  / Patent No. 5891695
5  / GENERAL INFORMATION:
6  / APPLICANT: Blanc, Veronique
7  / APPLICANT: Bianche, Francis
8  / APPLICANT: Crouzet, Joel
9  / APPLICANT: Jacques, Nathalie
10 / APPLICANT: Lacroix, Patricia
11 / APPLICANT: Thibaut, Denis
12 / APPLICANT: Zagorec, Monique
13 / APPLICANT: Debussche, Laurent
14 / APPLICANT: De Crecy-Lagard, Valerie
15 / TITLE OF INVENTION: Polypeptides Involved In The
16 / TITLE OF INVENTION: Biosyntheses Of Streptogramins, Nucleoside Sequences
17 / TITLE OF INVENTION: Coding For These Polypeptides And Their Use
18 / TITLES OF INVENTION: 43
19 / NUMBER OF SEQUENCES: 43
20 / CORRESPONDENCE ADDRESS:
21 / ADDRESSES: Fimegan, Henderson, Parabow, Garrett & Dunner
22 / STREET: 1300 I Street, N.W., Suite 700
23 / CITY: Washington
24 / STATE: D.C.
25 / COUNTRY: USA
26 / ZIP: 20005-3315
27 / COMPUTER READABLE FORM:
28 / MEDIUM TYPE: Floppy disk
29 / COMPUTER: IBM PC compatible
30 / OPERATING SYSTEM: PC-DOS/MS-DOS
31 / SOFTWARE: PatentIn Release #1.0, Version #1.30
32 / CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/403,852D
 FILING DATE: 10-MAY-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/FR 93/00923
 FILING DATE: 25-SEP-1993
 PRIOR APPLICATION DATA: FR 92/11441
 APPLICATION NUMBER: FR 92/11441
 FILING DATE: 25-SEP-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03806.0054-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4000
 TELEFAX: (202) 408-4400
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1833 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: S.pristinaespiralis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 103..1689
 US-08-403-852D-6

Query Match 4.7%; Score 78.6; DB 2; Length 1833;
 Best Local Similarity 45.4%; Pred. No. 4.8e-06;
 Matches 399; Conservative 0; Mismatches 474; Indels 6; Gaps 3;

414 GCTGCGGACACGGTGTGCGCTAATGAGCCGACCGCCGAGCGCGGTGTTCTACAGCGT 473
 408 GCTGCTGATGCTCGAGTGTGCTCTTCGCGCGCGGCTCGCTGCGCGCGCTTCCGCCA 467
 474 GAACGAGCGGAGCGCGGTGCTTTCACATCGCGGTGAGCGCGCGCGCGCTCTGAGG 533
 468 GGAATCGCGCACTATCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527
 534 GCTCATTTGATGTACCGCATCAACGAGAGTGCAGCTTCTGAGAGCGCGCTTCGCTGA 593
 528 GACCAACACCGCTGCGCGCTCATGAGAGTCTTCGAGAGAGAGAGCGCGCGCGCGAT 587
 594 CAGCTGAGCGCGCGCGCGCGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 653
 588 CGGCTGTGAGGAGCGCGCGCATCACTGCGCGCTTCGCGCGCGCGCGCGCGCGCGCG 647
 654 CCAAGAGCGGCGCACTTGACAAACAGAGCTGAGAGCAACAGAGTGTGAGCGCGCG 713
 648 CTTCTGACCACTTCTGTGAGGCGCTCATCTTCTGATCACTGCGCGCTCGCGCGCT 707
 714 GGGCACTGAGCG 773
 708 GGGCTGTGCGCGTGC - CCGCTGTGTGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 765
 774 CATTCTGTGAGCG 833
 766 CCGCACTGTGCGCGCGCGCGCGCTCTCAACCTCGGAGTGTGCGCGCGCGCGCGCATC 825
 834 CAGCTGCGCTTTCAGCGCAACAGCGCGCGCGCGAGCTGTGAGCGCTGTGCGCGCAAGT 893
 826 ATCTCTGCGCGCGCGAGACCGGCTGAGAGCGCGCGCG - AGGTCTCTGTGCGCGCGCG 884
 894 GAGCTGCGCAACCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCTTCTCTCGAGCGCGCG 953
 885 GGGCGCGCGCGCTGACCGGCTTCGTCGCTGAGAGAGTCAACAGAGAGAGAGAGAGAG 944
 954 GCCCGGAGAGGCTTCTTCTGTGAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCT 1013

Db 945 CATGAGCTTCTTACACCGACCGCGCTTCAACGAGCGCGCTCGCGGAGAGTGTCTCGTT 1004
 Qy 1014 GGCCTTCCGACATCAGTGTGTGAGCCCGCGCGTGTACGCGCGCAAGAGTGTGCGCGGA 1073
 Db 1005 CGGCAATGCGCGCTCGCTTCTCTGTCTCAACCGAGACCTCCAACTGTCTCGCTACGA 1064
 Qy 1074 CCCAAGCGCGCTGTGACCGCGAAGAGTACTGTGTGTGTGCGCGCGCGCGCGCGTGC 1133
 Db 1065 CCGCTGTGAGCGCGCGCTGTGCGGACCGCGCACTGTGTGAGATGTGTGCGCGTCAACT 1124
 Qy 1134 GAGCGCGCGCGAGCGG---CTCAGCTTCAAGTGTGCGCGCGCGCGCGAGAGTGTCTGTG 1190
 Db 1125 GCGCGCGCTCGCGCGGAAATCTTGTGCGCGCTTCGCGAGCGCGCGCGAGATGCGCTG 1184
 Qy 1191 CATCAACGCGCGTGTGCGCGCGCGCGCGCTGTGTGTGCTGACACCAAGAGCGCTGTG 1250
 Db 1185 CATGACATGTGTGCGCGCGCGCGCTGAGCGCGGTGTGCGCGCGAGATGCGCGCGAGCG 1244
 Qy 1251 CTTCTTGTGCG 1289
 Db 1245 CGGCTACGCGCGAGATGTGTGCGCGCGCGCTGTCTCAATGAG 1283

RESULT 10
 US-08-510-646B-6
 Sequence 6, Application US/08510646B
 Patent No. 6077699
 GENERAL INFORMATION:
 APPLICANT: Blanc, Veronique
 APPLICANT: Crouzet, Joel
 APPLICANT: Jacques, Nathalie
 APPLICANT: Lacroix, Patricia
 APPLICANT: Thibaut, Denis
 APPLICANT: Zagorec, Monique
 APPLICANT: Debussche, Laurent
 APPLICANT: De Crecy-Lagard, Valerie
 TITLE OF INVENTION: Polypeptides Involved In The
 TITLE OF INVENTION: Biosynthesis Of Streptogramin, Nucleotide Sequences
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESSES:
 ADDRESSER: Pimegan, Henderson, Farabow, Garrett & Dunner
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/510,646B
 FILING DATE: 03-AUG-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/403,852
 FILING DATE: 10-MAY-1995
 PRIOR APPLICATION DATA: PCT/FR 93/00923
 APPLICATION NUMBER: PCT/FR 93/00923
 FILING DATE: 25-SEP-1993
 PRIOR APPLICATION DATA: FR 92/11441
 APPLICATION NUMBER: FR 92/11441
 FILING DATE: 25-SEP-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03806.0054-01000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4000
 TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1833 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: S.pristinaespiralis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 103..1689
 US-08-510-646B-6

Query Match 4.7%; Score 78.6; DB 3; Length 1833;
 Best Local Similarity 45.4%; Pred. No. 4.8e-06;
 Matches 399; Conservative 0; Mismatches 474; Indels 6; Gaps 3;

414 GCTGCGCACACGCTGCTGCTACTGAGCGCCGACCGCGCGCGCTGTTACAGCGT 473
 408 GCTGCTGATGCTCGGACTGCTGCTTGGCGCGCGGCTGCGCGCTTGGCCCA 467
 474 GAACGACGCGGACCGGCTGCTTTCACCTGCGCGGCTGCGCGCGCGCGCTGAGG 533
 468 GACATCCGCGCCACATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGAGC 527
 534 GCTCATATGATGCTACGCGATACCGACGAGGTGAGCTTTCGAGAGCGCGCTTGTG 593
 528 GACCAACACCTCGCGCGCTGATCATGAGTCTTTCGACGACGACGACGCGCGCGCG 587
 594 CACGCTGACGCGCGCGCGCGCTGACGCGCGCGCGCTTACGCGCTGCGCGCGCG 653
 588 CGGCTGAGCGCGCGCGCGCGCGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 647
 654 CCAGACG 713
 648 CTTCTTCACACCTTCTGCTGCGCGCGCTCATTCTTCTGATCAACCTGCGCGCGCT 707
 714 GGGCCACCTGCG 773
 708 GGGCGCTGCTGCG 765
 774 CATTCGTCG 833
 766 CCGACCTGCTGCG 825
 834 CGACCTGCG 893
 826 ATCTCCG 884
 894 GGCCTGCG 953
 885 GGCCTGCG 944
 954 GCCCGCGCGAGACCTTCTGTCGAGGTGCGCGCGCGCGCGCGCGCGCGCGCG 1013
 945 CATGGCTTCTTCAACCGACCGCGCGCTTCAACGGCGCGCGCGCGCGAGTGTCTG 1004
 1014 GGCCTTGGCATCATGTCGTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1073
 1005 CGGCAATGCG 1064
 1074 CCCAAGCGCGCTGTCGACCGCGAAGATGTCGTCGTCGTCGTCGTCGTCGTCG 1133
 1065 GCGCTGCGACG 1124
 1134 GACCGCGCGCGACG 1190
 1125 GCG 1184
 1191 CATCAACG 1250

DB 1185 CATGACACTGCTGCGCGCGCGCGCTCAGACGCGGTGCGCGCGCGATCGCGCGCGCG 1244
 GY 1251 CTTCTTGGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1289
 DB 1245 CGGCTACGCGCGCGCATGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1283

RESULT 11
 US-09-231-818-6
 Sequence 6, Application US/09231818

Patent No. 6171846

GENERAL INFORMATION:

APPLICANT: Blanc, Veronique

APPLICANT: Blanc, Francis

APPLICANT: Crouzet, Joel

APPLICANT: Jacques, Nathalie

APPLICANT: Lacroix, Patricia

APPLICANT: Thibaut, Denis

APPLICANT: Zagorec, Monique

APPLICANT: Debussche, Laurent

APPLICANT: De Crecy-Lagard, Valerie

TITLE OF INVENTION: Polypeptides Involved In The

TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences

TITLE OF INVENTION: Coding For These Polypeptides And Their Use

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flunegan, Henderson, Farabow, Garrett & Dunner

STREET: 1300 I Street, N.W., Suite 700

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/231,818

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/403,852

FILING DATE: 10-MAY-1995

APPLICATION NUMBER: PCT/FR 93/00923

FILING DATE: 25-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 92/11441

FILING DATE: 25-SEP-1992

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146

REFERENCE/DOCKET NUMBER: 03806, 0054-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4000

TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1833 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: S.pristinaespiralis

FEATURE:

NAME/KEY: CDS

LOCATION: 103..1689

Query Match 4.7%; Score 78.6; DB 3; Length 1833;

LENGTH: 4257
 TYPE: DNA
 ORGANISM: HERPES VIRUS, TYPE 1
 US-09-259-821A-1

Query Match 4.6%; Score 77.6; DB 3; Length 4257;
 Best Local Similarity 43.3%; Pred. No. 7.4e-06;
 Matches 548; Conservative 0; Mismatches 694; Indels 24; Gaps 3;

8 AACAGGTGACCGGACCGCTTCCAGACCGCCAGCCCGGCGCTTCCGACACCGGCG 67
 Db 1853 ACGAGGGGGTCCCGCGCTGCGCGCGCGACCGGCGCGCGCGCGCGCGCGGT 1912
 QY 68 CGTGTGCG 127
 Db 1913 ACG 1972
 QY 128 CGCAGCG 187
 Db 1973 CGGCG 2032
 QY 188 ACAAGCTTCTGCAATGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 247
 Db 2033 GCG 2092
 QY 248 GCGTGTGCG 307
 Db 2093 CGTGTGCG 2152
 QY 308 CGCAGCG 367
 Db 2153 CG 2212
 QY 368 TCACG 427
 Db 2213 CGCAGCG 2269
 QY 428 TGTGTGCG 487
 Db 2270 TGTGTGCG 2329
 QY 488 CGTGTGCG 547
 Db 2330 CG 2389
 QY 548 ACG 589
 Db 2390 CG 2449
 QY 590 CTGACAGCG 649
 Db 2450 TGTGTGCG 2509
 QY 650 GCAGCG 709
 Db 2510 CG 2569
 QY 710 AGTGTGCG 769
 Db 2570 CG 2629
 QY 770 CG 829
 Db 2630 GCTGTGCG 2689
 QY 830 AGCG 889
 Db 2690 AGCG 2749
 QY 890 AAGTGTGCG 949
 Db 2750 CGGTGTGCG 2809

QY 950 TGTGTGCG 1009
 Db 2810 AGCG 2869
 QY 1010 CGGTGTGCG 1069
 Db 2870 CG 2929
 QY 1070 CGCAGCG 1129
 Db 2930 CGGTGTGCG 2986
 QY 1130 TGTGTGCG 3046
 Db 2987 CG 3106
 QY 1190 GCATGAGCG 1249
 Db 3047 GCG 3106
 QY 1250 CCTTCT 1255
 Db 3107 CGGCGCT 3112

RESULT 15

US-08-843-659-1
 Sequence 1, Application us/08843659

Patent No. 6218103

GENERAL INFORMATION:

APPLICANT: Leopardi, Roarlio

APPLICANT: Roizman, Bernard

TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND IC4 AS

NUMBER OF INVENTION: INHIBITORS OF APOPTOSIS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSES: Arnold, White & Durkee

STREET: P. O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: United States

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/843,659

FILING DATE: Concurrently Herewith

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

REFERENCE/DOCKET NUMBER: ARSB:519

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4257 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-843-659-1

Query Match 4.6%; Score 77.6; DB 3; Length 4257;
 Best Local Similarity 43.3%; Pred. No. 7.4e-06;
 Matches 548; Conservative 0; Mismatches 694; Indels 24; Gaps 3;

8 AACAGGTGACCGGACCGCTTCCAGACCGCCAGCCCGGCGCTTCCGACACCGGCG 67
 Db 1853 ACGAGGGGGTCCCGCGCTGCGCGCGCGACCGGCGCGCGCGCGCGCGGT 1912

QY 68 CAGTCTGCGGCCCCGAGCGAGCCCGGTCCTGAGGAGCGCGCTTCCACG 127
 Db 1913 ACCGCGCGCGGAGATCTGCGCGCGCTGAGGCGCGCTGCGCGCGCGCTTCCCCCG 1972
 QY 128 CGCAGCGCAAGGAGAGAGATGAGGCTGAGCGGCGCATCGCGCGCGCGCAAGGCGCA 187
 Db 1973 CGGAGGCGAGAGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2032
 QY 188 ACAGCTTCTGCAATGAGGCTCACTTCAAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 247
 Db 2033 GCGCGCAGGCG 2092
 QY 248 GAGTGGCGCTGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 307
 Db 2093 CGCTGGCGGAGGCGCTTGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2152
 QY 308 CGCAGCATTCGCTGCTCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
 Db 2153 CCGCGAGCG 2212
 QY 368 TCACG 427
 Db 2213 CCGAGCG 2269
 QY 428 TGGTGGCGCTACTGAGCG 487
 Db 2270 TGGTGGCTCATGCG 2329
 QY 488 CGGTGCTCTTCACTGCG 547
 Db 2330 CCG 2389
 QY 548 ACCGCGAT-----CACCGAGAGGTCAGCTTCTGAGAGCGCGCTTTCG 589
 Db 2390 CCG 2449
 QY 590 CTGACAGCGTGAAGCG 649
 Db 2450 TGGCG 2509
 QY 650 GCGAGCGAGAGCG 709
 Db 2510 CCG 2569
 QY 710 AGCTGGGCGACCTGAGCG 769
 Db 2570 CGCGCGGAGGCG 2629
 QY 770 CCGCGCATTCGTCG 829
 Db 2630 GCTCGAGAGCG 2689
 QY 830 AGGCGGAGCTGCGCGCTTCCAGCGMAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 889
 Db 2690 AGCG 2749
 QY 890 AAGTGGCGCTGCGCAAGCG 949
 Db 2750 CCGTGGCGCGTTCG 2809
 QY 950 TGCAGCGCGCGCGAGAGCTTTCGTGAGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1009
 Db 2810 AGCG 2869
 QY 1010 GCGTGGCGCTTGGGATCACTGTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1069
 Db 2870 CCG 2929
 QY 1070 CCGAGCGCGAGCGCGCTGCTGACCGCAAGAGTACTGAGTGTGAGCGCGCGCGCGCGCGCG 1129
 Db 2930 CCTCTAGTGTGATCCCG 2986

QY 1130 TGGCGAGCGGCGGAGAGCGGCTGAGCTTCAAGCTGAGCGCGCGCGCGCGCGCGCGCG 1189
 Db 2987 CCG 3046
 QY 1190 GCATCAAGGCGGCTGCG 1249
 Db 3047 GGGCG 3106
 QY 1250 CTTTCT 1255
 Db 3107 CGGCGCT 3112

Search completed: January 20, 2004, 18:00:10
 Job time : 139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 20, 2004, 14:33:57 : Search time 6062 Seconds

(without alignments)
11303.798 Million cell updates/sec

Title: US-09-808-387-21

Perfect score: 1675

Sequence: 1 atgggcaacacgggtgcacgg.....ctacagcgcatagcctacgac 1675

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_ov:*

22: em_ov:*

23: em_pac:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrc:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	724.8	43.3	189479	9	AC027309	AC027309 Homo sapi
2	724.8	43.3	193960	9	AC022217	AC022217 Homo sapi
3	724.8	43.3	326249	2	AC034210	AC034210 Homo sapi
C 4	499.4	29.8	219695	10	AC122252	AC122252 Mus muscu
C 5	449.2	26.8	238906	2	AC095473	AC095473 Rattus no
C 6	449.2	26.8	277331	2	AC110151	AC110151 Rattus no
7	409.2	24.4	2207	6	E26374	E26374 Neurelized
8	409.2	24.4	2207	9	U87864	U87864 Homo sapien
9	409.2	24.4	3634	9	AF029729	AF029729 Homo sapi
10	407.6	24.3	4011	9	BC026336	BC026336 Homo sapi
11	399.6	23.9	2445	10	MMNBURALI	Y15160 Mus musculu
12	396.4	23.7	2210	10	AF400063	AF400063 Mus muscu
13	393.2	23.5	3630	10	AF401228	AF401228 Mus muscu
14	392.2	23.4	1641	6	E26375	E26375 Neurelized
C 15	323.4	19.3	277331	2	AC110151	AC110151 Rattus no
C 16	199.6	11.9	3176	5	AF419159	AF419159 Xenopus 1
C 17	186.8	11.2	168481	2	AC025094	AC025094 Homo sapi
C 18	186.8	11.2	203262	9	HSBA416N2	AL121929 Human DNA
19	183.4	10.9	1005	10	MM0271921	AJ271921 Mus muscu
C 20	178.6	10.7	178832	2	AC129799	AC129799 Rattus no
C 21	178.6	10.7	186929	10	AC090657	AC090657 Mus muscu
C 22	178.6	10.7	191825	2	AC132288	AC132288 Mus muscu
C 23	178.6	10.7	226140	2	AC097415	AC097415 Rattus no
C 24	178.6	10.7	233713	2	AC133034	AC133034 Rattus no
C 25	178.6	10.7	282915	2	AC098165	AC098165 Rattus no
26	177.2	10.6	193627	2	AC091427	AC091427 Mus muscu
C 27	170.8	10.2	110000	2	AC12401_1	Continuation (2 of
28	166.4	9.9	1704	9	BC012317	BC012317 Homo sapi
29	166.4	9.9	1733	9	IR2004990	AL889981 Homo sapi
C 30	165.4	9.9	236486	9	AC021188	AC021188 Homo sapi
C 31	158.4	9.5	2207	3	DVNEUR82	U12592 Drosophila
C 32	138.4	8.3	168337	2	AC118956	AC118956 Rattus no
C 33	135.8	8.1	2539	10	AF321278	AF321278 Mus muscu
C 34	134.8	8.0	224371	2	AC132456	AC132456 Mus muscu
35	128.6	7.7	1198	10	MM0271920	AJ271920 Mus muscu
36	127	7.6	2656	3	AY051987	AY051987 Drosophila
37	127	7.6	3396	3	BT003772	BT003772 Drosophila
C 38	127	7.6	123579	2	AC017792	AC017792 Drosophila
C 39	127	7.6	176036	3	AC009258	AC009258 Drosophila
C 40	127	7.6	306651	3	AB003681	AB003681 Drosophila
41	126.6	7.6	3542	3	S62597	S62597 neu-neural1
42	125.4	7.5	3999	3	DROC3HC4ZF	L12218 Drosophila
43	125.4	7.5	3999	3	S62583	S62583 neu-neural1
C 44	119	7.1	46965	2	AC006183	AC006183 Homo sapi
45	115.8	6.9	110000	2	AL831721_0	AL831721 Danto rer

ALIGNMENTS

RESULT 1
AC027309 189479 bp DNA linear PRI 30-JUL-2002
LOCUS Homo sapiens chromosome 5 clone CTB-7988, complete sequence.
DEFINITION
AC027309
ACCESSION
VERSION
AC027309.3 GI:22004171
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (Bases 1 to 189479)
AUTHORS
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE
Direct Submmission
JOURNAL
Unpublished

REFERENCE	2 (bases 1 to 189479)
AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submission
JOURNAL	Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE	3 (bases 1 to 189479)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Submitted (30-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On Jul 30, 2002 this sequence version replaced gi:7712023. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov
FEATURES	Finishing Completed at Stanford Human Genome Center www.hgsc.stanford.edu Quality: Phrap Quality >=40 99.9% of Sequence; Estimated Total Number of Errors is 0.2. Location/Qualifiers
SOURCE	1. 189479 /organism="Homo sapiens" /mol_type="Genomic DNA" /db_xref="taxon:9606" /chromosome="5" /clone="CTB-79B"
BASE COUNT	47394 a 47381 c 47862 g 46842 t
ORIGIN	
Query Match	43.3%; Score 724.8; DB 9; Length 189479;
Best Local Similarity	98.4%; Pred. No. 9.3e-70;
Matches 732; Conservative	0; Mismatches 12; Indels 0; Gaps 0;
OY	577 GAGAGCGCCTTGTGCTGACACGCCTGATGCCCGCCGCTTCAGCCAGGCCCGCTTCAAGCCC
DB	128961 GAGAGCGCCTTGTGCTGACACGCCTGATGCCCGCCGCTTCAGCCAGGCCCGCTTCAAGCCC
OY	637 TGCCCTGGCCCCAGACGCCAGACGCGCGGCCCACTTGACAACAACAGAGCTCGAAGAACA
DB	129021 TGCCCTGGCCCCAGACGCCAGACGCGCGGCCCACTTGACAACAACAGAGCTCGAAGAACA
OY	697 CAGGTGTGTGCCCAAGCTGGGACCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
DB	129081 CAGGTGTGTGCCCAAGCTGGGACCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
OY	757 GACGCGCGGCGCGCGCCATTCCGTGTGGGCCCCCGGTGAGCGCGCGCGCGCGCGCGCTGTGTG
DB	129141 GACGCGCGGCGCGCGCCATTCCGTGTGGGCCCCCGGTGAGCGCGCGCGCGCGCGCTGTGTG
OY	817 CCGGCGCTACTGGAGAGGCCGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
DB	129201 CCGGCGCTACTGGAGAGGCCGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
OY	877 TCGGCGCACCGCAAAATGTGCTGTGCGAACCGCGGCGCGGACCGGCGCGGACCGCTGTCTTC
DB	129261 TCGGCGCACCGCAAAATGTGCTGTGCGAACCGCGGCGCGGACCGGCGCGGACCGCTGTCTTC
OY	937 TCAGAGCGCCCGCTGTGCGCGCGCGCGCGGACCTTTGTGTGAAGTGTGGACCGTCCGCGGACTG
DB	129321 TCAGAGCGCCCGCTGTGCGCGCGCGCGGACCTTTGTGTGAAGTGTGGACCGTCCGCGGACTG
OY	997 GCGGCGCCCGCGCGCTGTGCGCTTTCGCGCATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
DB	129381 GCGGCGCCCGCGCGCTGTGCGCTTTCGCGCATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
OY	1057 AACGAGCTGCCCCGCGAACCCAGACCGGCGGTGTGAACCGCAAAAGATGTAAGTGGGTGTGGGCG
DB	129441 AACGAGCTGCCCCGCGAACCCAGACCGGCGGTGTGAACCGCAAAAGATGTAAGTGGGTGTGGGCG
OY	1117 CGCGCGCGGCGCGGTGTGCGGACGCGCGGACCGGCTTCAGCTTCACTGTGTGTGTGTGTGTGTGT
DB	129501 CGCGCGCGGCGCGGTGTGCGGACGCGCGGACCGGCTTCAGCTTCACTGTGTGTGTGTGTGTGTGT
OY	1177 GACGTGCTCTTGGGCAATCAAGGCGTCCGCGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGT

```

Db      12961  GACGTGCTCCGGGACATCAACGGGCGTCCGGCGGCGCCGCTGTGTGCGTGCACACACAG 129620
Oy      1237   CAGCGCGCTTGGGCTTTTGTGGCGGTGGCGGCGCGTGCAGCGGCAAGTGTGTCCTTC 1296
        |||
Db      129621  CAGCGCGCTTGGGCTTTTGTGGCGGTGGCGGCGCGTGCAGCGGCAAGTGTGTCCTTC 129680
Oy      1297   GGTACCGCTGCAGTGCAGCGCTGCG 1320
        |||
Db      129681  GGTACGTCCTCCGCGCGCGCGCGTCCG 129704

RESULT 2
AC022217
LOCUS      AC022217      193960 bp      DNA      linear      PRI 09-MAY-2001
DEFINITION Homo sapiens BAC clone RP11-779018 from 5, complete sequence.
ACCESSION  AC022217
VERSION     AC022217.5  GI:14010943
KEYWORDS   HTG.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens (human)
            Bufo vulgaris, Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 193960)
            Sulston, J.B. and Waterston, R.
            Toward a complete human genome sequence
            Genome Res. 8 (11), 1097-1108 (1998)
JOURNAL     99063792
MEDLINE     9847074
PUBMED      9847074
REFERENCE   2 (bases 1 to 193960)
            Kruchowski, S., Stoneking, T., Hawking, M., Strommat, C. and Le, T.
            The sequence of Homo sapiens BAC clone RP11-779018
            Unpublished
JOURNAL     3 (bases 1 to 193960)
            Waterston, R.H.
            Direct Submissio
            Submitted (09-MAY-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            4 (bases 1 to 193960)
            Waterston, R.
            Direct Submissio
            Submitted (09-MAY-2001) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On May 9, 2001 this sequence version replaced gi:11245628.
JOURNAL     -----
            Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: MUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: saplens@watson.wustl.edu
            Summary Statistics
            Center project name: H_NH0779018
            -----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McHersom, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:

```

The RPc1-11 human BAC library was made from the blood of one male donor, as described by Oseegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tatematsu, M., Catenean, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION: The clone sequenced to the right is RP11-536N17. Actual start of this clone is at base position 1 of RP11-779018, actual end is at base position 193960 of RP11-779018.

There are polymorphic base pair differences in the overlap between the clone RP11-779018 and RP11-536N17.

Location/Qualifiers

FEATURES

Bouyce

repeat_region	/organism="Homo sapiens"
repeat_region	/mol_type="genomic DNA"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="5"
repeat_region	/map="5"
repeat_region	/clone_lib="RPCL-11"
repeat_region	2..98
repeat_region	/rpt_family="L2"
repeat_region	100..343
repeat_region	/rpt_family="MIR"
repeat_region	563..635
repeat_region	/rpt_family="L1"
repeat_region	659..822
repeat_region	/rpt_family="L1"
repeat_region	852..1212
repeat_region	/rpt_family="MaLR"
repeat_region	1218..1349
repeat_region	/rpt_family="L1"
repeat_region	1423..1921
repeat_region	/note="similar to EST AA429140 (NTD:g2111915) zw51c10.r11"
repeat_region	1582..1869
repeat_region	/rpt_family="MBR1_type"
repeat_region	1157..2089
repeat_region	/note="similar to EST AV706616 (NTD:g10725881)"
repeat_region	1982..2117
repeat_region	/rpt_family="L2"
repeat_region	2145..2209
repeat_region	/note="similar to EST AV706616 (NTD:g10725881)"
repeat_region	2265..2560
repeat_region	/rpt_family="Alu"
repeat_region	2627..2704
repeat_region	/rpt_family="L2"
repeat_region	2715..2947
repeat_region	/rpt_family="MIR"
repeat_region	3252..4496
repeat_region	/note="CPG island (%GC=72.5, o/e=1.00, #CPGs=177)"
repeat_region	3390..3543
repeat_region	/note="similar to EST A1571301 (NTD:g4534675) tnc4a606.x1"
repeat_region	3402..3539
repeat_region	/note="similar to EST AW489411 (NTD:g7059681)"
repeat_region	3403..3539
repeat_region	/note="similar to EST BE649566 (NTD:g9975390)"
repeat_region	3562..3742
repeat_region	/note="similar to EST BE649566 (NTD:g9975390)"
repeat_region	3562..3736
repeat_region	/note="similar to EST AW489411 (NTD:g7059681)"
repeat_region	3755..4116
repeat_region	/note="similar to EST BE097995 (NTD:g8488889)"
repeat_region	3756..4116
repeat_region	/note="similar to EST AMS29983 (NTD:g7172397)"
repeat_region	3756..4110
repeat_region	/note="similar to EST AM049442 (NTD:g5909971)"
repeat_region	3756..4108
repeat_region	/note="similar to EST A1579407 (NTD:g4563783)"

misc_feature	3763..4110	/note="similar to	BST	AA850748	(NID:g2938288) "
misc_feature	3775..4116	/note="similar to	BST	AA851139	(NID:g2938679) "
repeat_region	4278..4407	/rpt_family="MBR1_type"			
repeat_region	4947..5094	/rpt_family="MBR1_type"			
repeat_region	5071..5192	/rpt_family="MBR1_type"			
misc_feature	5296..5611	/note="similar to	BST	AA873871	(NID:g2966007) ch79d08.sa"
repeat_region	5500..5879	/rpt_family="MBLR"			
misc_feature	6143..6268	/note="similar to	BST	AA850748	(NID:g2938288) "
misc_feature	6143..6257	/note="similar to	BST	AA049442	(NID:g5909971) "
misc_feature	6652..6849	/note="similar to	BST	BB863961	(NID:g10384525) "
misc_feature	6653..6821	/note="similar to	BST	AA850748	(NID:g2938288) "
misc_feature	7789..8086	/note="similar to	BST	BB062153	(NID:g8406803) "
misc_feature	7844..8018	/note="similar to	BST	BB695120	(NID:g10082280) "
misc_feature	7978..8567	/note="similar to	BST	AL041833	(NID:g5421179) "
misc_feature	8407..9088	/note="similar to	BST	AA862418	(NID:g7958116) "
misc_feature	8656..9301	/note="similar to	BST	BB888251	(NID:g10344366) "
misc_feature	8788..9267	/note="similar to	BST	A1276718	(NID:g3888992) q175b12.x1"
misc_feature	8847..9359	/note="similar to	BST	AA248434	(NID:g6591427) "
misc_feature	8983..9557	/note="similar to	BST	AA377155	(NID:g6881818) "
misc_feature	9201..9353	/note="similar to	BST	AA091410	(NID:g1636873) "
misc_feature	9359..10116	/note="similar to	BST	AA683237	(NID:g10285100) "
repeat_region	9374..9538	/rpt_family="Tc2"			
misc_feature	9500..9879	/note="similar to	BST	T08307	(NID:g389335) "
repeat_region	9540..9668	/rpt_family="MBR1_type"			
misc_feature	9771..10430	/note="similar to	BST	AA953119	(NID:g8142802) "
misc_feature	9771..10137	/note="similar to	BST	AA348548	(NID:g2001787) "
misc_feature	9790..10190	/note="similar to	BST	R13230	(NID:g766306) yf74e08.r1"
misc_feature	9821..10194	/note="similar to	BST	AA392404	(NID:g2046417) zt7b09.r1"
misc_feature	10016..10655	/note="similar to	BST	BB783751	(NID:g10204949) "
misc_feature	10088..10929	/note="similar to	BST	BB274435	(NID:g9149376) "
misc_feature	10267..10421	/note="similar to	BST	AA091410	(NID:g1636873) "
misc_feature	10344..11196	/note="similar to	BST	BB908856	(NID:g10403854) "
repeat_region	10346..10407	/rpt_family="L2"			
misc_feature	10358..10844				

Query Match	43.3%	Score 724.8;	DB 9;	Length 193960;
Best Local Similarity	98.4%;	Pred. No. 9.2e-70;		
Matches 732; Conservative	0;	Mismatches 12;	Indels 0;	Gaps 0;

517 GAGAGCGCTTTCGTCGACACGCTACGCGCGCGGCTCAAGCCAGAGCCCGCTTACGCCCC 636


```

Db      267612 AACGAGCTGCCCCGACCCAGACGCGCTCTGACCGCAAGATGATGAGTGGCG 267553
QY      1117 CCGGCGGGGCGCCCTGCGGAGCGCGGAGACCGCTCAAGTTCAACGCTGCGGCGCGCG 1176
Db      267552 CCGGCGGGGCGCCCTGCGGAGCGCGGAGACCGCTCAAGTTCAACGCTGCGGCGCG 267493
QY      1177 GACGTGCTCTGGGCAATCAACGAGGCGTCCGCGGCGCGCTGCTGAGCTGACACACG 1236
Db      267492 GACGTGCTCTGGGCAATCAACGAGGCGTCCGCGGCGCGCTGCTGAGCTGACACACG 267433
QY      1237 CAGGCGCTCTGGGCGCTTCTTTCGCGGCGCGCGCGCTGCGGCGCGCTGCTTCTTC 1296
Db      267432 CAGGCGCTCTGGGCGCTTCTTTCGCGGCGCGCGCGCTGCGGCGCGCTGCTTCTTC 267373
QY      1297 GGTACCTGAGTCCAGCCTTGGC 1320
Db      267372 GGTGAGTCCCGGCGCGCGCTGGC 267349

RESULT 4
AC122252/c      219695 bp      DNA      1linear      ROD 29-MAY-2003
LOCUS      Mus musculus chromosome 17 clone RP23-170E7, complete sequence.
DEFINITION      AC122252
ACCESSION      AC122252.4 GI:31126889
VERSION      HTG
KEYWORDS      Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 219695)
AUTHORS      Wilson, R.K.
TITLE      The sequence of Mus musculus clone
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 219695)
AUTHORS      McPherson, J.D. and Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE      3 (bases 1 to 219695)
AUTHORS      McPherson, J.D. and Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (18-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE      4 (bases 1 to 219695)
AUTHORS      Wilson, R.K.
TITLE      Direct Submission
JOURNAL      Submitted (29-MAY-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT      On May 29, 2003 this sequence version replaced gi:29029447.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BA0170E07
-----
FEATURES
Source      1. 219695
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-170E7"
/clone="RP23-170E7"

BASE COUNT      54661 a 54739 c 54797 g 55498 t

Query Match      29.8%; Score 499.4; DB 10; Length 219695;
Best Local Similarity 82.3%; Pred. No. 1,3e-45;
Matches 597; Conservative 0; Mismatches 101; Indels 27; Gaps 1;

```

```

QY      575 TGAAGAGCGCTTCTGCTGACAGCTGACCGCGCGCTTCAAGCGAGCGCTTCAAGCG 634
Db      204239 TAGAAAGACACTTTCCTGACAGCTGACCGCGCGCTTCAAGCGAGCGCTTCAAGCG 204180
QY      635 CTGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 694
Db      204179 CTGCGCGCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 204120
QY      695 ACCAGGTGAGCGCAAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 754
Db      204119 ACCAGGTGAGCGCAAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 204076
QY      755 CCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 814
Db      204075 -----GAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 204027
QY      815 CCGCGCGCGCTACTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 874
Db      204026 CCGCGCGCGCTACTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203967
QY      875 TGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 934
Db      203966 TGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203907
QY      935 TCTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 994
Db      203906 TCTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203847
QY      995 TGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1054
Db      203846 TGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203787
QY      1055 CCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1114
Db      203786 CATCCAGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203727
QY      1115 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1174
Db      203726 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 203667
QY      1175 GCGAGGTGCTCTGAGGCAATCAACGAGGCGTCCGCGCGCGCGCGCGCGCGCGCG 1234
Db      203666 GCGAGGTGCTCTGAGGCAATCAACGAGGCGTCCGCGCGCGCGCGCGCGCGCG 203607
QY      1235 CCGAGGCGCGCTGAGGCGCTTCTTTCGCGGCGCGCGCGCGCGCGCGCGCGCG 1294
Db      203606 CACAGGCGCGCTGAGGCGCTTCTTTCGCGGCGCGCGCGCGCGCGCGCGCGCG 203547
QY      1295 TCGGT 1239
Db      203546 TCGGT 203542

RESULT 5
AC095473/c      238906 bp      DNA      1linear      HTG 09-MAY-2003
LOCUS      Rattus norvegicus clone CH230-713, WORKING DRAFT SEQUENCE.
DEFINITION      AC095473
ACCESSION      AC095473
VERSION      AC095473.6 GI:30467680
KEYWORDS      HTG, HTGS_PHASE2, HTGS_DRAFT, HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE      1 (bases 1 to 238906)
AUTHORS      Muzny, D., Matile, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayegbe, A., Ayodeji, M., Baca, B., Baden, H.,
Baldwin, D., Bandaranatke, D., Barber, M., Barnstead, M., Benham, F.,
Biswal, K., Blair, J., Blankenburg, K., Blythe, P., Brown, M.,
Bryant, N., Bunney, C., Burch, F., Burrell, K., Calderon, B.,

```

Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Dublin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Faller, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huh, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., Kling, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzuhwa, L., Loulsegod, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangun, A., Mangun, B., Mague, P., Martin, K., Martin, R., Martinez, B., McWhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Mogensavljevic, A., Miner, G., Miranda, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeleneh, O., Okunnu, G., Olarnunagsoon, A., Pal, S., Parks, K., Paatenkash, S., Paul, H., Perez, A., Perez, L., Plamnick, C., Plopper, F., Polndexter, A., Popovic, D., Primus, E., Pu, L., -L., Puzo, M., Quirco, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Rella, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X. -Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Syatek, A., Tabors, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Direct Submission
2 (bases 1 to 238906)
Morley, K.C.

Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 238906)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:2491075.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

```

Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information -----
Center project name: GCKH
Center clone name: CH230-713
----- Summary Statistics -----
Assembly program: Atlas;
Consensus quality: 226530 bases at least Q40
Consensus quality: 226594 bases at least Q30
Consensus quality: 231141 bases at least Q20
Estimated insert size: 236988; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Gendbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
  1 238906: contig of 238906 bp in length.
FEATURES
    source
        1..238906
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-713"
misc_feature
    1..1595
    /note="wgs_contig"
BASE COUNT
    57521 a 59144 c 57758 g 57899 t 6584 others
ORIGIN
Query Match
    26.8%; Score 449.2; DB 2; Length 238906;
    Best Local Similarity 87.8%; Pred. No. 3e-40;
    Matches 490; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
    Db
        84 CCCCAGCGAAGCCCGGCTCTGGGCGAAGGCGCCGCTTCAAGCGCAAGGCCAAAGGCAA 143
        81091 CCGGAGAGGCGAGCGGCTCTGGGCGAAGGCGCCAGCTTCAAGCGCAAGGCCAAAGGCAA 81032
    Oy
        144 GAACGTGGCGCTGAGAGCGCACTCGGCGCGGCGCAACAGCGCGAAAGCTTCGCAATG 203
        81031 GAACGTGGCGCTGAGAGCGCACTCGGCGCGGCGCAACAGCGCGAAAGCTTCGCAACG 80977
    Oy
        204 CGTCAAGTTACGCAAGGCGCCATCCGCGCTGTACAGAGAGTTCGCGCTGCGCTGGTGC 263
        80971 AGTCACTTACGCAAGGCGCCATTCGCGCTGTACAGAGAGTTCGCGCTGCGCTGGTGC 80912
    Db
        264 CGTGGCGCTGAGAGCGGCGCGCTGCGCTTCGCGCTTCAAGCGCGAAGATCGTGCCT 323
        80911 TGTGCGTCTGAGAGCGGCGCGCGCTTCGCGCTTCAAGCGCGAAGATCGTGCCT 80855
    Oy
        324 CATGAGGCGCCAGAGCACTCCCAAGTACGCTTCGCGAGACTTGGATCAACCGCGCGGCGTA 383
        80851 CATGAGGCGCCAGAGATATCCCAAGTACGCTTCGCGAGACTTGGATCAACAGACTTGAATA 80795
    Oy
        384 CTGGGCGCAAGGCACTGCGCGAGAACTGCGCTGCGCAACGATCTGGCTTAAGTGC 443
        80791 CTGGGCGCAAGGCGCTGCGCGAGAACTGCGCTGCGCAACGATCTGGCTTAAGTGC 80733
    Db
        444 CGACCGCGACGCGCGCGTGTCTTACAGCGTGAACGAGCGGAGCCGATGCTTTCACTG 503
        80731 TGAATCGTCAAGGTCGCGCTTCTTATAGTGTCAATGATGCGAACAAGTGTCTGTTCACTG 80672
    Oy
        504 CGGCGTGGCGGTGGGCGGCGCGCTCTGGGCGCTCATTTGATGTTCAGGGCATCAGACGA 563

```

Db	Accession	Version	Source	Organism	Reference	Authors
Db	80671	CGCGGTGGCCCTGGAGAGGCCCACTCTGGGCACTATGCACCTCTATGTCATCAGCAGCA	80611	GATCAGCTGCTGGGTAG	80594	
Db	564	GATCAGCTTCTGAGAG	581			
Db	80611	GATCAGCTGCTGGGTAG	80594			
RESULT 6	AC110151	277331 bp	DNA	linear	HTG 22-SEP-2002	
LOCUS	AC110151					
DEFINITION	Rattus norvegicus clone CH230-47B19, *** SEQUENCING IN PROGRESS					
ACCESSION	AC110151					
VERSION	AC110151.4	GI:23195250				
SOURCE	HTG, HTGS, PHASE1, HTGS DRAFT, HTGS_ENRICHED.					
ORGANISM	Rattus norvegicus (Norway rat)					
	Rattus norvegicus					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;					
	Rattus.					
	1 (bases 1 to 277331)					
	Murphy, D. Marie, Metzger, M. Lee, Abramson, S., Adams, C., Alder, J.,					
	Allen, C., Allen, H., Albrooke, S., Amin, A., Anguiano, D.,					
	Antalech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,					
	Baldwin, D., Bandaranaike, D., Barber, M., Barnes, M., Beahmed, F.,					
	Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,					
	Bryant, N., Buhay, C., Burck, P., Burrell, K., Calderon, E.,					
	Cadenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,					
	Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,					
	Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., Souza, L.,					
	Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,					
	Delgado, O., Denson, S., Desmo, C., Ding, Y., Dinh, H., Divya, K.,					
	Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,					
	Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,					
	Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,					
	Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Gatta, M.,					
	Georgopoulos, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,					
	Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K.,					
	Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J.,					
	Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,					
	Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,					
	Jackson, L., Jacod, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,					
	Kapachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, J.,					
	Kows, C., Kraft, C. L., Leboy, H., Levan, J., Lewis, L., Li, Z., Liu, J.,					
	Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,					
	Lorensheva, L., Loutseg, H., Lozano, R. J., Lu, X., Ma, J.,					
	Maneshwar, M., Mahindratne, M., Mamoud, M., Malloy, K., Mangum, A.,					
	Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,					
	Mathew, S., McLeod, M. P., McNally, T. Z., Meenen, B.,					
	Milosavljevic, A., Miner, G., Minja, E., Montemayor, J.,					
	Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,					
	Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,					
	Nwoketeme, O., Okwono, G., Olarinmaga, A., Pal, S., Parks, K.,					
	Paetern, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,					
	Plommer, F., Poindester, A., Popovic, D., Primus, E., Pu, L.,					
	Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,					
	Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,					
	Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,					
	Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,					
	Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smaj, D.,					
	Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J.,					
	Steinle, M., Strong, R., Sutton, A., Svetek, A., Taber, P., Taylor, C.,					
	Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trojes, Z., Uman, K.,					
	Valas, R., Vera, V., Villanar, D., Waldron, L., Walker, B., Wang, J.,					
	Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,					

```

REFERENCE
AUTHORS      2 (bases 1 to 277331)
TITLE        Worley K.C.
JOURNAL      Direct Submission
             Submitted (10-FEB-2002) Human Genome Sequencing Center, Department
             of Molecular and Human Genetics, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
REFERENCE
AUTHORS      3 (bases 1 to 277331)
TITLE        Rat Genome Sequencing Consortium.
JOURNAL      Direct Submission
             Submitted (22-SEP-2002) Human Genome Sequencing Center, Department
             of Molecular and Human Genetics, Baylor College of Medicine, One
             Baylor Plaza, Houston, TX 77030, USA
COMMENT
             On Sep 19, 2002 this sequence version replaced g1:21743237.
             The sequence in this assembly is a combination of BAC based reads
             and whole genome shotgun sequencing reads assembled using Arlas
             (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
             sequence may extend beyond the ends of the clone and there may be
             contigs that consist entirely of whole genome shotgun sequence
             reads. Both end sequences and whole genome shotgun sequence only
             contigs will be indicated in the feature table.
             -----
             Genome Center
             Center: Baylor College of Medicine
             Center code: BCM
             Web site: http://www.hgsc.bcm.tmc.edu/
             Contact: hgsc-help@bcm.tmc.edu
             -----
             Project Information
             Center project name: GRPU
             Center clone name: CH230-47B19
             -----
             Summary Statistics
             Assembly program: Phrap; version 0.990329
             Consensus quality: 231941 bases at least Q40
             Consensus quality: 236540 bases at least Q30
             Consensus quality: 239843 bases at least Q20
             Estimated insert size: 252686; sum-of-contigs estimation
             Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
             -----
             * NOTE: Estimated insert size may differ from sequence length
             * (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
             * NOTE: This sequence may represent more than one clone.
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 3 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *
             *
             1      13219: contig of 13219 bp in length
             *
             *      13220      13319: gap of unknown length
             *
             *      13320      266558: contig of 253233 bp in length
             *
             *      266559      266658: gap of unknown length
             *
             *      266659      277331: contig of 10673 bp in length.
             *
             Location/Qualifiers
             1..277331
             /organism="Rattus norvegicus"
             /mol_type="genomic DNA"
             /db_xref="taxon:10116"
             /clone="CH230-47B19"
             416..1276
             /note="clone boundary
             clone_end:5p6
             site:BCORI
             end_sequence:BH356625"
             36181..37518
             /note="wgs contig"
             41065..42752
             /note="wgs_contig"
             83582..86052
             /note="wgs_contig"
             106338..107848
             /note="wgs contig"
             137921..139155
             /note="wgs_contig"

```

misc_feature	complement (245796..246299)
	/note="clone_boundary clone_end:T7 site:ECORI end_sequence:BH356621" 265182..266558 /note="wgs_end_extension clone_end:T7"
misc_feature	
BASE COUNT	60560 a 61304 c 60133 g 60025 t 35309 others
ORIGIN	
Query Match	26.8%; Score 449.2; DB 2; Length 277331;
Best Local Similarity	87.8%; Pred. No. 2,9e-40;
Matches 490; Conservative	0; Mismatches 68; Indels 0; Gaps 0;
Db	24 CTTGCCAGACCCGAGCCCAACCGGCGCGCTCTTGAGCCACCGGCGCTGCGGCCCGG 83
Db	38624 CCCCACAGACTCCAGCCCTCTTGACCGCTCTTGAGCCACCGGCGCTTGCTACCGGCCGG 38683
Db	84 CCCCAGAGCAGCGCCGGCTCTTGAGCGAGGCGCGCGCTTTCACAGCCGACCAAGACCA 143
Db	38684 CCCCAGAGGAGCAGCGGCTCTTGAGCGAGGCGCGCCACGCTTTCACAGCGGACGACCA 38743
Db	144 GAAGTGGGAGCTGAGCGGCGCATCTGCGCGCGGCGCCACAGCGGCGCAAGCTTTCAGATG 203
Db	38744 GAAGTGGGAGCTGAGCGGCGCATCTGCGCGCGGCGCCACAGCGGCGCAAGCTTTCAGATG 38803
Db	204 CGTACGCTTTCACGAGCGGCGCCATCCGCTGTACAGACAGTGCAGCTGCGCTGTGCG 263
Db	38804 AGTACCTTTCACGAGCGGCGCCATCCGCTGTACAGACAGTGCAGCTGCGCTGTGCG 38863
Db	264 CGTGCCTCTTGAGCGGCGCGCTGCGCTTTCACAGCGGCGCAAGTCTGCTGCT 323
Db	38864 TGTGCTCTTGAGCGGCGCGCTGCGCTTTCACAGCGGCGCAAGTCTGCTGCT 38922
Db	324 CATGAGGCGCCAGGACATCCCAAGTACGCTGCGCGGACCTGCTACAGCGGCGCGGCTTA 383
Db	38924 CATGAGGCGCCAGGACATCCCAAGTACGCTGCGCGGACCTGCTACAGCGGCGCGGCTTA 38983
Db	384 CTGGGCGCAAGCATCTGCCGAGAACCTTGCGCTGCAGACAGTGTGCGCTTACTGCGC 443
Db	38984 CTGGGCGCAAGCGCTGCCGAGAACCTTGCGCTGCAGACAGTGTGCGCTTACTGCGC 39043
Db	444 CGACCGCCACGCGCGCTGTTCACAGCTGTAAAGCAGCGGAGCGGCTGCTTTCACATG 503
Db	39044 TGATCTGTCACGTGCCTTCTTATGTGTCAATGATGGCGAACAGTCTGTTCCATG 39103
Db	504 CGGCTGCGCGCTGCGCGCGCGCTCTTGAGCGCTCATGATGTTCACGCGATTCACGACGA 563
Db	39104 CGGCTGCGCGCTGCGCGCGCGCATCTTGAGCGCTCATGATGTTCACGCGATTCACGACGA 39163
Db	564 GGTGAGCTTCTGAGAG 581
Db	39164 GGTGAGCTTCTGAGAG 39181

COMMENT	OS	PN	JP	UP	1991	37257-A/1
PR	MOTOMI NAKADA, HIDEO	NAKAMURA, MITSUHIRO	YOSHIDA, HIDEYUKI	SAYA		
PC	C12N15/09, C07K14/47, C07K16/18, C12P21/02, C12Q1/68, G01N33/53//					
PC	(C12N15/09, C12R1:91), (C12P21/02, C12N1:19), C12N15/00, (C12N15/00, PC					
C12R1:91)						
CC	Strandedness: Double;					
CC	Topology: Linear;					
FT	Key					
FT	source					
FT	Location/Qualifiers					
FT	1..2207					
	/organism="Unidentified".					
	/organism="Unidentified"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:32644"					
BASE COUNT	340 a	872 c	673 g	322 t		
ORIGIN						
Query Match	24.4%;	Score 409.2;	DB 6;	Length 2207;		
Best Local Similarity	56.6%;	Pred. No. 2.7-35;				
Matches	896;	Conservative	0;	Mismatches 628;	Indels 60;	Gaps 5;
QY	99	GATCTTGAGGCGGCGCGCTTTCACCGCGCAGCCCAAGGCAAGAACTGCGGCTTGA	158			
DB	581	GCTCCGACGACGCGCGCTGCTTTCACCGCGCAGCCCAAGGCTTCCAGATCTCATGA	640			
QY	159	CGGCACTCGCGCGCGCGCGCACCGCGCGCAAGCTTTCGCAATGCGTCACTGACGCA	218			
DB	641	CCTCAGCACAAGGCTTTCAGAGGCGCGCGCGCTTTCGCAACCCCATCTTCAGCA	700			
QY	219	GCGGCGCATCGCGCTGATCAGCAGAGGCGCGCTGCGCTGAGTGGCGCGCGCTGCTG	278			
DB	701	CGCGCGGCTCTCATCTACAGCAGGCTGAGCTGAAATCACCAGAGCAGTGTGCTG	760			
QY	279	GAGCGGCGCGCTGCGCTTTCACCGCGCAGATCTGCTCATGAGCGCCAGGA	338			
DB	761	GAGCGGCGCGCTGCGCTTTCACCGCGCAGATCTGCTCATGAGCGCCAGGA	820			
QY	339	CATCCCAAGTACGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	398			
DB	821	GCTGCCAAGTACGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	880			
QY	399	GCGCGAAGCTGCG	458			
DB	881	GCGCGAAGCTTTCAGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	940			
QY	459	CGTGTTCACG	518			
DB	941	TGTTCACG	1000			
QY	519	CGGCGCGCTGCGCGCGCTCATTTATGCTTACGCGCATCACGAGAGTCAAGCTTTCGA	578			
DB	1001	CGAAGCGCTGCG	1060			
QY	579	GAGCGCGCTGCG	638			
DB	1061	TAGGAGCTG-----TGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1111			
QY	639	CGTCCG	698			
DB	1112	GCG	1166			
QY	699	GAGTGTGCG	758			
DB	1161	-----AGCTTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1210			
QY	759	CG	818			

Db 1211 GGCAGCCGCTGCCCCCATCCGCAAGACTCACTCAACTGACGACAGCCGCGCTGCC 1270
 Oy GGGGCTACTGAGAGCCGACTGCGCTTCACAGCAACAGCGAGGCGCG---ACGTAGAGCT 875
 Db 1271 GGGGAGCTTCAAGCGGACCTGCTTTCAGAGCCCTGCGGCGCGGCGGACGTCGCTT 1330
 Oy 876 GTCCGCGCAGCGCAAGTGGCTGCGACCGCGGCGGACGCGGCGGCGGCTGCTT 935
 Db 1331 CCTGACGAGGACGAGCGGTGGCGCGGTGAGAGACGGGCGGACGAGCGCGCTGCTT 1390
 Oy 936 CTCCGAGCGCGCGCTGCGGCGCGGAGAGCTTCTTGTGAGAGTGGCGCTCGCGGCT 995
 Db 1391 CACCAAGCGCGCGCGGTGCGGCGGAGACCACTTCTGTCAGAGTCAAGCGCGCTGCGG 1450
 Oy 996 GCGCGCGCGCGCGCGCTGCGGCTTGGGATCACTGTCGAGCGCGGCGGCGGCTGAGCC 1055
 Db 1451 GCGCGCGCGCGCGCGCTGCTGCTGCGGCTCACTGTCAGCGCGGCGGCGGCTGCGG 1510
 Oy 1056 CAACGAGCTGCGCGCGGCGGACCGGACGCGCTGCTGACGCGCAAGAGTACTGGTGTG 1115
 Db 1511 GCGCGAGCTGCTTTCAGCGCTTGAAGCGCGGCTGAGCGCAAGAGTCTGGGCGCTG 1570
 Oy 1116 GCGCGCGCGCGCGCGGCGGAGCGGCGGCGGCGGCTGAGCTTCAAGCTTCCGCGG 1175
 Db 1571 CCGCGTGGCGCGCGCGCTGCAAGCGGCGGCACTCTGGGCGCTGTCAGCGCGAG 1630
 Oy 1176 CGAGCTGCTGCGGCGGATCAACGCGGCGGCGGCGGCGGCGGCTGTCGAGCGAGC 1235
 Db 1631 CGAGCTGACCTTCAAGCGGCAAGTGGCGGCGGCGGCGGCGGCGGCTGTCGAGCG 1690
 Oy 1236 GCAAGCGCTGCTGCGGCTTCTTCCGCGGCGGCGGCGGCGGCGGCGGCTGCTCT 1295
 Db 1691 GCAAGCGCTGCTGCGGCTTCTTCCGCGGCGGCGGCGGCGGCGGCGGCTGCTCT 1750
 Oy 1296 CGGTACCTGCGAGTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTCT 1346
 Db 1751 CTCCACTATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTCT 1810
 Oy 1347 -----CGGCTCCAGGAGCATAGTATGATTCAGATTAAGCTTCAAGTCAACAGTC 1397
 Db 1811 AACTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1870
 Oy 1398 CTCTCTGCGATCTGAGTCACTCCCTGCGGAGCGGCGGCGGCGGCGGCGGCTG 1457
 Db 1871 TGGCCCTCTGCGGAGCTGCTGCTGCGGAGCGGCGGCGGCGGCGGCGGCGGCTG 1930
 Oy 1458 GTCCCGCTGCTTCTCCCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1517
 Db 1931 GTCCCGAGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1981
 Oy 1518 CTTCGATGCGGAGTGGAGACAGGATCACTCAAGTGGAGCAATGTCGTCAGCAG 1577
 Db 1982 CTATGAACAGCGGAGTGGAGACAGGATCACTCAAGTGGAGCAATGTCGTCAGC 2041
 Oy 1578 CTGCGGCGCTGCGGCTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1637
 Db 2042 CTGTGGCGCTGCGGCTCAAGAGGCTGCAAGCGGCGGCGGCGGCGGCGGCTG 2101
 Oy 1638 CAAGAGCTGATTAAGATTTACAG 1661
 Db 2102 CAAGAGCATCATCAAGAGCTTACCG 2125

RESULT 8

U87864 2207 bp mRNA linear pri 27-MAY-1998
 LOCUS U87864 Homo sapiens neutralized homolog (h-neutralized) mRNA, complete cds.
 DEFINITION U87864
 ACCESSION U87864.1 GI:3157990
 VERSION U87864.1
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 2207)
 AUTHORS Nakamura,H., Yoshida,M., Teulki,H., Ito,K., Uno,M., Nakeo,M., Oke,K., Tada,M., Kochi,M., Kuratsu,J., Ushio,Y. and Sawa,H.
 TITLE Identification of a human homolog of the Drosophila neutralized gene within the 10q25.1 malignant astrocytoma deletion region
 JOURNAL Oncogene 16 (8), 1009-1019 (1998)
 MEDLINE 98178690
 PUBMED 9519875
 REFERENCE 2 (bases 1 to 2207)
 AUTHORS Yoshida,M., Nakamura,H., Teulki,H., Makino,K. and Sawa,H.
 TITLE Direct Submission
 JOURNAL Submitted (30-JUN-1997) Tumor Genetics and Biology, Kumamoto University School of Medicine, Honjo 2-2-1, Kumamoto 860, Japan

FEATURES
 source
 1..2207
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="brain"
 /dev_stage="fetal"
 1..2207
 /gene="h-neutralized"
 411..2135
 /note="ring finger protein, similar to Drosophila neutralized, encoded by Genbank Accession Number L12218"
 /codon_start=1
 /product="neutralized homolog"
 /protein_id="AAC17474.1"
 /db_xref="GI:3157991"
 /translation="MGNNPSSIPSLPRNGPBRPGRPHQNDKISIGPPVTSRCH KQKCPALVPSGGIAPPLPFPPTKQSLIMDSHAKVQAFCAITPSNPVLI YBOVRLKTKKQCGSALRLGFTSPSRJHPSLPPLADPLVAVGLTGVOLDS FANGNIIAFVNDKKGRVFRHINDSAVNLPSGVRADPLMALVDVYGLTGVOLDS ELVLPDCLRPSPALRPSLRARRADARLSVSLCDLVPGADDBAALPAGCPDPN SLNQHQBALPQIDGRLRPAALAAAGATVRLDQTVARVHGRDBALVPTSPRVY ARTIFVKTRSGARPAALSGVTCDDGTLPADLPFPSPALVDKRFVAVCRVGP LHSQDIDLVNNAQBELHSHNGAAGQQLCVDSQPMPLFGHGTITQRLIAGTI LAERGIPLSPSPALPTSPALSLRSLDPLSTGSGSGLSSAGTAPNSVLSIPS PVTGLQGWSDBCITCYEHAVDVYITCGHNCICVACGLRLKKALHACCPICRRPIKD IIKYTRSS"
 1971..2090
 /gene="h-neutralized"
 /note="ring finger motif"

BASE COUNT 340 a 872 c 673 g 322 t

ORIGIN
 Query Match 24.4% Score 409.2, DB 9, Length 2207;
 Best Local Similarity 56.6% Pred. No. 2.7e-35;
 Matches 896, Conservative 0, Mismatches 628, Indels 60, Gaps 5;

Oy 99 GGTCTGGGAGAGGCGCGGCTTCCAGCGGCGGCAAGGCAAGAACTGCGGCTGGA 158
 Db 581 GCTCCACACAGCGCGGCTTTCACCGGCAACAGAGGCTCCAGATCTCATGGA 640
 Oy 159 CGGCACTCGCGCGCGGCAACAGCTTCTGCAATGCGCTCACTGACGCA 218
 Db 641 CCTCAGCCACAAGCTGTCAAGAGCGAGCGAGCTTCTGCAAGCGCATCACTTACGAA 700
 Oy 219 GCGGCCATCGGCTGTACAGACAGTGGCGGTGCGGCTGAGTGGCGGCGGCTG 278
 Db 701 CCGCCGGTCTCATCTACAGACAGTCAAGCTGAAATACCAAGAGCAAGTGTCTG 760
 Oy 279 GAGCGGCGGCTGCGGCTTCCAGCGGCGGCAAGATCGTCTCATGAGCGCCAGGA 338
 Db 761 GAGCGGCGGCTTCCAGCGGCTTCCAGCGGCAAGACCGGTCCGATTCACCTGATC 820
 Oy 339 CATCCCAAGTACGCTTCCGCGGCAAGCTGATCAAGCGGCGGCTACTGAGCGCAAGGCACT 398
 Db 821 GCTGCCAAGTACGCTTCCGCGGCAAGCTGATCAAGCGGCGGCTTCTGAGCGCAAGGCGCT 880
 Oy 399 GCCCGAAGACTGCGGCTGCGGCAAGCTGATGCTACTGAGGCGGCAAGCGGCGGCG 458


```

Db      881  GCTGAGGAGTTGGCCAAATGAGGGCAACATCATTCGATTCTGGGTGACAAAGAGGCGG 940
Qy      459  CGTGTCTTAAGCGGTGAACAGCGGAGCGCGTCTCTTCACTGCGCGGTGCGGTGAG 518
Db      941  TGTCTTCAACCGCATCAAGCACTCGGTGTATGCTGTTCTTCAAGCGGGGTCCGACGGC 1000
Qy      519  CGGCGGCTCTGGGCGCTCATATGATCTAGGCAATCAAGGAGAGTGAAGTCTTGTGA 578
Db      1001  CCAACCGGCTCTGGGCGCTGTGTGACGTCTACAGGCTCAAGGAGGAGTCAAGCTGTGA 1060
Qy      579  GAGCGGCTTGTGAGACAGCTGACGCGCGGCGCTCAAGCAGGCGCGCTTCAAGCGCTG 638
Db      1061  TACCGAGCTGG-----TGCCTCCGAGCTGTCTGCGGCGCGCTCTTCAACCGGCT 1111
Qy      639  CTTGCGCGCCAGAGCAAGACGCGGCACTTTCAGCAACAGAGCTCGAAGAACCA 698
Db      1112  GCGGCGGCGGTGCTGCGGCGGAGGCGAGCAAGCGCGGCTCTCGGTG----- 1160
Qy      699  GGTGTGGCCAGCTGGGCACTGTGGCGCTGGGCGCGCGCGCGCGCGCGCGCGCGCG 758
Db      1161  -----AGCTTATGCACTCAACGTGCGGCGCGAGCGGCGAGCGGCGCGGCGC 1210
Qy      759  CGCGCGGCGCGCGCAATTCGTCGCGGCGCGGCGCGGCGCGCGCGCGCGCGCGCG 818
Db      1211  GCGCGCGGCGTCCGCTCCGAGAACTACTCACTCACTCGAGACAGCGCGCGCTGCC 1270
Qy      819  GCGGCTACTGAGAGCGGCACTGTGCTTCAAGCAACAGCGGCGCGG---ACGTGAGCT 875
Db      1271  GGGGCAAGCTGAGAGGAGCACTGTGCTTCAAGCGCGCGCGCGCGCGCGCGCGCAT 1330
Qy      876  GTGCGCGGACCGCAAGTGTGCTGCGGCAAGCGGCGCGCGCGCGCGCGCGCGCGCT 935
Db      1331  CCTCGAGAGCAAGCGTGTGCGCGCGGTGAGAGCAAGCGCGCGGAGAGCGCGCTCT 1390
Qy      936  CTCGAGAGCGCGCGCTGCGCGCGCGCGCGGAGGCTCTTCTGAGAGGTGGCGCGCG 995
Db      1391  CACGAGCGGCGCGGTGCGGTGCGCGAGCAATCTTCTGTAAGGTCAAGCGCTCGGTG 1450
Qy      996  GCGGCGCGCGCGCGCGCTGTGCTTTCGATCAAGTGTGAGCAAGCGCGCGGTGTAG 1055
Db      1451  CGGCGCGCGCGCGCGCGGTGTCTGTTGCGGCTCAAGAGTGTGAGCGCGCGCGCG 1510
Qy      1056  CAAAGAGCTGCGCGCGAGCAAGCGGCTGTGAGCGGAAAGATCTGGGTGTGCGC 1115
Db      1511  GCGCGAGCTGCTTTCAGGCGCTGAGCGCGGTGAGCGGAGAAATCTGGGCGCGGTG 1570
Qy      1116  GCGCGCGCGCGCGCGGTGCGAGCGCGCGCGCGCGCTTCAAGCTTCAAGCTGCGCG 1175
Db      1571  CCGCGCGCGCGCGCGCGGTGAGAGCGGCAATCTGAGGCTGTGATCAAGCGCAAG 1630
Qy      1176  CGAGCTGCTCTGGGCACTCAAGCGGCGTCCGCGCGCGCGCGCTGTGCTGCAACAC 1235
Db      1631  CGAGCTGCACTCAAGCAATAGGCGCGCGCGCGCGCGGTGAGTGTGAGCGGCTG 1690
Qy      1236  GCAGCGCTGTGGAGCTTCTTCCGCGGTGCGCGCGCGCGCGGTGCGGCGAGTGTGCT 1295
Db      1691  GCAGCGGCTTGTGATGCTTTCGCGCTGCAAGGAGCATCAAGAGATTCGCACTCGG 1750
Qy      1296  CGGTACCTGCAATCAAGCGCTGCAAGCAAGCTCAATCAAGGAGCGCTTCAAG----- 1346
Db      1751  CTCACATCTGTGAGCGAGGAGGTATCCGCTACCTCCCTGCTCTTCCCTTCAGCGC 1810
Qy      1347  -----CGGTCCAGAGAGATAGTATTCAGATATGACCTTCAAGTCAACAGTC 1397
Db      1811  AACCTGCGCCAGTGCCTGTGGAGCGCGCGCTGTGAGCCCTGTGTAGCAAGTGAAGCTC 1870
Qy      1398  CTCCTGGAGATCTGATGATCTCTGTGAGCGGCGCGCGCGCGCGCTGAGCGCGCGGT 1457
Db      1871  TGGCGCTCTGTGAGCTCTGTGTGAGGAGCGCGCGCAATTCGCAATGAGCTGCGCGA 1930
Qy      1458  GTCCCCGCTGTTCTCCCAACCGAGCGGAGGAGCATCAAGATGGCGAGTGAAGTGTG 1517

```

```

Db      1931  GTGCGAGTGAACCCAGGTCTGTGGGCGAGT-----GAGGATGAGTGCACATTTG 1981
Qy      1518  CTTCAATGCGAGGTGAGACAGGCTCATCTACAGCTGTGACACATGTCCTGTGCGCAG 1577
Db      1982  CTATGAAACAGCGGTGAGACAGGTATCTACACATGTGCGCAATGTGCTTGTACAGC 2041
Qy      1578  CTGCGGCTGCGGCTCAAGGAGAGGCGCGGCTGTGCGCGCGCGCGCGCGCGCGCG 1637
Db      2042  CTGTGCGCTGCGGCTCAAGAGGCTGTGACAGGCTGTGCGCGCGCGCGCGCGCAT 2101
Qy      1638  CAGAGCTCATTAAGATCTACAG 1661
Db      2102  CAAGACATCATCAAGACTACCG 2125

RESULT 9
AF029729      3634 bp      mRNA      linear      PRI 05-JAN-1999
LOCUS      Homo sapiens neutralized mRNA, complete cds.
DEFINITION
ACCESSION      AF029728
VERSION      AF029729.1 GI:4103927
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
            Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 3634)
AUTHORS      Prinos, P., Kilpatrick, M.W. and Tsipouras, P.
TITL      Direct Submision
JOURNAL      Submitted (10-OCT-1997) Pediatrics, UCONN Health Center, 263
            Farmington Ave., Farmington, CT 06030, USA

FEATURES
    source
        1..3634
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="10"
            /map="110g24-q25"
            62..1786
            /codon_start=1
            /product="neutralized"
            /protein_id="AAD01887.1"
            /db_xref="GI:4103928"
            /translation="MGNNPSSIPLPRGNSPAPRGHNTLNDISGAPPVTSRCHH
            KQNCAPVLPSGGLPATPLLPHTGSGQILDLSKAVRQASFCNATTPSRPILI
            YQVRLKIKTKCCGSGALRLGTSKDSRPHIDSLPKTACPDLVQSGFWALALPB
            PANBNIIAFWDKGVRPHRINDSAVMKFPFGVTRADPLMLVDVYGLTRGVLDS
            BLVLPCLRP8FTLALRPSLRBADLISLSDLVNFGADGDAAPAGCPIDN
            SLNDSHRALPQLDGLRPHALRAGAVRILDSOTVAVHGRDRLALVTSRPPRV
            ARTIPKVTBGGAPRCAISFGVTTCDPTLPRADLPSPBALVDRKPMVCRVGP
            LNSGDLGLVNADGELHSHNGAAMQCLVDASGPLMMLRELHGTTOIRLIGSTI
            LABRGIPSLPSPASTPTSPALSGSLSDPLSTGSGSPGSSAGGTAPSPVLSBS
            PVPPIGQMSDECTTCYBHAVDVIVTTCGMCLCYAGRLKALHACPICRPIKD
            ILYKRS"

BASE COUNT      614 a      1192 c      1076 g      752 t
ORIGIN
Query Match      24.4%; Score 409.2; DB 9; Length 3634;
Best Local Similarity 56.6%; Pred. No. 2.3e-35;
Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

```


279 GAGCGCGCGCTGCGCTTCCGCTTCAACGCGGACGATCCGTGCTCATAGAGGCCAGGA 338
 412 GAGCGCGCGCGCTGCGCTTCCGCTTCAACGCGGACGATCCGTGCTCATAGAGGCCAGGA 471
 339 CATCCCGCAAGTACGCTGCGCGGACCTGCTCAACGCGCGCGCTGCTGCGCGCGGACCT 398
 472 GCTGCCCAAGTACGCTGCGCGGACCTGCTCAACGCGCGCGCTGCTGCGCGCGGACCT 531
 399 GCCCGAAGACCTGCGCGGACCTGCTCAACGCGCGCGCTGCTGCGCGCGGACCTGCGCG 458
 532 GCTCGAAGAGTTCGCAATGAGGGGCAATCATGCGATTCCTGGGAGGCAAGAGGGGCG 591
 459 GGTGTTTCAAGCTGGAACGAGCGGAGCGGCTGCTTCACTGCGCGCGGCTGCGCGCGG 518
 592 TGTCTTCAACGCGCACTCAAGCTGCTTCAAGCTGCTTCAAGCTGCTTCAAGCTGCTTCA 651
 519 GCGCGCGCGCTGCGCGGACCTGCTTCAAGCTGCTTCAAGCTGCTTCAAGCTGCTTCA 578
 652 GAGCGCGCGCTGCGCGGACCTGCTTCAAGCTGCTTCAAGCTGCTTCAAGCTGCTTCA 711
 579 GAGCGCGCTTCCGCTGCAACGCTGACCGCGCGCGCTGCTGACGCGCGCGCTTCAAGCG 638
 712 TAGCGAGCTGG-----TGTCCCGGACGTGCTGCGCGCGCGCTTCAAGCGCGCT 762
 639 CTGCGCGCGCGGACG 698
 763 GCGCGCGCGCTGCTGCG 811
 699 GGTGTTGCGCAAGCTGCGCGGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 758
 812 -----ACCTATGCGACCTCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 861
 759 GCG 818
 862 GCG 921
 819 GCGCGCTTCAAGCTGCGCGGACCTGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCGCG 875
 922 GCGCGAGCTGCGCGGACCTGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 981
 876 GTCG 935
 962 CCTCGAAGACGAGCGGCG 1041
 936 CTCCGAGCG 995
 1042 CACGAGCG 1101
 996 GCG 1055
 1102 GCG 1161
 1056 CAAGAGCTGCG 1115
 1162 GCGCGAGCTGCG 1221
 1116 GCG 1175
 1222 CCG 1281
 1176 GCGAGCTGCTTCAAGCTGCGCGGACCTGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCG 1235
 1282 GAGCTGCGACCTCAAGCTGCGCGGACCTGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCG 1341
 1236 GAGGCGCGCTGCGCGGCTTCAAGCTGCGCGGACCTGCTTCAAGCTGCGCGCGCGCGCG 1295
 1342 GAGGCGCGCTTCAAGCTGCGCGGACCTGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCG 1401
 1296 CGGTACCTTCAAGCTGCGCGGACCTGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCGCG 1348
 1402 CTCGACTATCTGCGCGGACCTGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1461
 1349 -----GCTCCGAGAGGATGATTCAGATTAAGCTTCAAGCTGCGCGCGCGCGCGCG 1397

Db 1462 AACCTCGCCAGTGCCTTGGGAGCGCGCTGCTGAGCCCTTCTGCTCAAGACGTGAGCTC 1521
 Gt 1398 CTCCTCGGACATCTGAGTATCTCTGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1457
 Db 1522 TGGCCCTCTGAGTGTCTGCTGAGGAGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1581
 Gt 1458 GTCCCGGTGTTCTCCCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1517
 Db 1582 GTCCCGGTGAGCG 1632
 Gt 1518 CTTCGATGCGGAGGAGGAGCAAGCTGCTTCAAGCTGCTTCAAGCTGCTTCAAGCTGCTTCA 1577
 Db 1633 CTATGAGACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1692
 Gt 1578 CTGCGCGCTGCGGCTCAAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1637
 Db 1693 CTGAGCGCTGCGGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1752
 Gt 1638 CAAGAGCTCATTAAGTCTACAG 1661
 Db 1753 CAAGAGCATCATCAAGACTACCG 1776

RESULT 10
 BC026336
 LOCUS
 DEFINITION Homo sapiens, neuralized-1 like (Drosophila), clone MGC:26482
 IMAG:4812302, mRNA, complete cds.
 ACCESSION BC026336
 VERSION BC026336.1 GI:20070954
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 4011)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbe-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shireki Tohhyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson Mark) mcdpaxll.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IPAX Plate: 32 Row: 3 Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4758799.
 FEATURES
 source
 1. 4011
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="LOCUSID:9148"
 /db_xref="taxon:9606"
 /clone="MGC:26482 IMAG:4812302"
 /tissue_type="Brain, hippocampus"
 /clone_id="NIH MGC_95"
 /lab_host="DH10B"
 /note="Vector: pBluescript"

CDS

98..1822
 /codon_start=1
 /product="neuralized-like (Drosophila)"
 /protein_id="AAH26336.1"
 /db_xref="GI:20070955"
 /translation="MGNNPSSIPSLPRGNPSRAPRPHQNLKDISGPPVTSRCH
 KKHCPAVLPSSGLPAALIFPHPTKGSOLIMDSHKAYRQASPCNALTFSPRLVLI
 YKQVRLKTKKQCCNMGALRLGTSQDPSIHDSIPKACPLVVSQSGFMAALBE
 FANEGNIIAFWVDRKGRVFRINDSAVMLEPFSVTRADPLMALVDYGLTRVQLDS
 BLVLPDLRPSFTALRPSLRREADARLSVSLCDLNPFGADGDAAPASCPIDN
 SINSOSRALPAQDLDRFHALPAGAHVYIIDEQTVARVEHDERALVTSRPARV
 AETIPKVTSGSGARPGALPSGVTCDPGLPADLPSPBALVDRKBPVAVRVPV
 LMSGLDILGVNADGELHSHNGAAGMOLCDASQPIPMLEPILHGTGRIILGSTI
 LAERGIPLSPSPASTPSPALSGSLDPLSTGSGPLSGSAGTAPSPISLES
 PVTPLGQMSDECTICYEHAVDVITYCGHMCVACGLKRLKALHACPTCKRPYK
 IIKYRS"

BASE COUNT 667 a 1303 c 1208 g 833 t

Query Match 24.3%; Score 407.6; DB 9; Length 4011;
 Best Local Similarity 56.5%; Pred. No. 3.3e-35;
 Matches 895; Conservative 0; Mismatches 629; Indels 60; Gaps 5;

QY 99 GGCCTGGGGGAGAGCGCGGCTTCCACGCGAGGCGCAAGAGCGTGGCGCTGA 158
 Db 268 GCTCCAGCGCGCGCGCTGCTTCCACCGCACACCAAGAGGCTCCAGATCTCATGGA 327
 QY 159 CGGCACTCGCGCGCGCGCAACGCGCGCAACGCTTGTGATAGCGTCACTTACGCA 218
 Db 328 CTTACAGCAACAGGCTGTCAAGAGGCGCGAGCTTGTGACGCGCATCTTACAGCA 387
 QY 219 GCGGCGCATCGGCTGTACAGAGAGTGGCGCTGCGCTGTGGCGCGCGCGCTG 278
 Db 388 CGCGCGCGCTCATCTACAGAGAGTCAAGTCAAGATCAACAGAGAGTGTGCTG 447
 QY 279 GAGCGGCGCGCTGCGCTTCAACGCGCGCAACGCTGCTGTGAGCGCGCGCA 338
 Db 448 GAGCGGCGCGCTGCGCTTCAACGCGCGCAACGCTGCTGTGAGCGCGCGCA 507
 QY 339 CATCCCAAGTACGCTGCGCGCACTGTGTACGCGCGCGCGCTACTGCGCAAGGCACT 398
 Db 508 GCTGCCAAGTACGCTGCGCGCGCACTGTGTGTCCAGAGTGGCTTGTGGCGCAAGGCGCT 567
 QY 399 GCGCGAGACCTGCGCGCTGCGCAACGCGTGTGCTTGTGGCGCGCGCGCGCG 458
 Db 568 GCGTGAAGATTGCGCATGAGGCGCAACATCATGCTTGTGGGTGAGCAAGAGGCGCG 627
 QY 459 CGGTCTTACAGGCTGAGAGCGCGCGAGCGGCTTTCACGCGCGCGCGCGCG 518
 Db 628 TGTCTTTCACCGATCAACGATCTGCTGTATGTGTGTCTTTCAGCGGCGCGCGCGCG 687
 QY 519 CGGCGCGCTTGTGGCGCTCATTTGATGTCTACGCGCATCACGAGCGAGTGTGAGCTTGTGA 578
 Db 688 CGACCGGCTTGTGGCGCTTGTGTGAGCTGTACGCGCTACCGCGGCGCGTCCAGGCTGTGA 747
 QY 579 GAGCGCGCTTGTGTGAGCGCTGAGCGCGCGCGCTTGTGAGCGCGCGCGCTTGTGAGCGCTG 638
 Db 748 TAGCGAGCTG-----TGTCCCGGAGCTGTCTGCGCGCGCGCTTGTGAGCGCGCT 798
 QY 639 CTGCGCGCGCGAGCGAGCGCGCGCACTTGTGAGCAACGAGCTGTGAGAAACGA 698
 Db 799 GCGCGCGCGCTGCTGCGCGCGAGCGAGCGAGCGCGCGCTTGTGAG-----847
 QY 699 GGTGTGTGCGCAAGCTGTGCGCACTGTGCGCGCGCGCGCGCGCGCGCGCGCA 758
 Db 848 -----AGCTATGTGCGCTCAACGAGTGTGCGCGCGCGCGAGCGAGCGCGCGCG 897
 QY 759 GCGCGCGCGCGCGCGCATTTCCGTGCGCGCGCGCGCGCGCGCGCGCGCGCTGTGCGCG 818
 Db 898 GCGCGCGCGCGCGCGCGCATTTCCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 957
 QY 819 GCGCGCTGTGAGCGCGAGCTTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 875

Db 958 GCGGAGCTGAGCGGAGCTGCTTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCAT 1017
 QY 876 GTTCGCGGACCGCAAGTGGCTGTGCGCACCGCGCGCGCGCGCGCGCGCGCGCTT 935
 Db 1018 CTTGACGAGCAGACGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 1077
 QY 936 CTTGAGCT 995
 Db 1078 CACGACCGCT 1137
 QY 996 GCG 1055
 Db 1138 CG 1197
 QY 1056 CAGCGAGCTGCG 1115
 Db 1198 GCGCGAGCTGCG 1257
 QY 1116 GCG 1175
 Db 1258 CG 1317
 QY 1176 CGACGCTCTCTGCGCGCATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
 Db 1318 CGAGCTGCACTCAGCGCAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
 QY 1236 GCGAGGCTCTGCGCGCTTCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1295
 Db 1378 GCGCGCGCTTGTGATCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
 QY 1296 CGGTACCGCTGAGCG 1348
 Db 1438 CTTCACTATCTGTGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
 QY 1349 -----GCTCCAGAGCAGATGATGATGATGATGATGATGATGATGATGATG 1397
 Db 1498 AACCTGCGCGAGCG 1557
 QY 1398 CTTCTGCGCATGTGATCATCTGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCG 1457
 Db 1558 TGGCGCTGTGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1617
 QY 1458 GTCGCCGCTTCCCGCACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1517
 Db 1618 GTCCAGAGTACCGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1668
 QY 1518 CTTGATGAGCGAGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1577
 Db 1669 CTATGAAACAGCGGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1728
 QY 1578 CTTGCGCGCTGCGCGCTCAAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1637
 Db 1729 CTTGCGCGCTGCGCGCTCAAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1788
 QY 1638 CAGGAGCTCATTAAGTATCAG 1661
 Db 1789 CAGGAGCTCATTAAGTATCAG 1812

RESULT 11
 MNEURALI 2445 bp mRNA linear ROD 07-JAN-2000
 LOCUS MNEURALI 2445 bp mRNA linear ROD 07-JAN-2000
 DEFINITION Mus musculus mRNA homologous to protein encoded by Drosophila
 neutralized gene.
 ACCESSION Y15160
 VERSION Y15160.1 GI:6688678
 KEYWORDS neutralized gene.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 Pavlopoulos, E., Prinos, P., Klapirick, M., Tsipouras, P. and

JOURNAL Maechonas, N.K.
 REFERENCE Unpublished
 2
 AUTHORS Moschonas, N.K.
 TITLE Direct Submission
 JOURNAL Submitted (17-OCT-1997) N.K. Moschonas, Dept of Biology, University of Crete and IMBB-FORTH, P.O.Box 1527, Heraklion 711 10, Crete, Greece

REMARK Revised by [3]
 3 (bases 1 to 2445)
 AUTHORS Moschonas, N.K.
 TITLE Direct Submission
 JOURNAL Submitted (10-FEB-1998) N.K. Moschonas, Dept of Biology, University of Crete and IMBB-FORTH, P.O.Box 1527, Heraklion 711 10, Crete, Greece

REMARK Revised by author 23-MAR-1998
 FEATURES Location/Qualifiers

1..2445
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="12A3"
 /tissue_type="brain"
 /clone_1fb="E14.5"
 /dev_stage="embryonic"
 1..1725
 /codon_start=1
 /product="neuralized-like protein"
 /protein_id="CA65238.1"
 /db_xref="GI:668679"
 /translation="MGNFSSVSLQGNPSPASRHPQNLDSIGSGFPVPSRCHH
 KOHCPPTLSGGGLPATPLFPHPTKGSQILMDLSHKAVKQASPCNAITFSNRPVL
 YEOVRLKTKKCCWGALRIGFTSKDSRIHPSLPRYACDLYSOGAPAKLPER
 PANENITIAFWDKGGRVYRINESAMTFSGVTPARIMLVNVCITGVQILDS
 ELVDPDLRPSFTALKRPSLCEADEARLSLDCIVANPGADGDAPRACPTLPQ
 SLNSQHSRALPAQLDGLRFLARRRAHRLIDEQTVARVERDERALVFTSRPLRV
 AETIPIKVTRSGGARPGALSGVTTCDPQTLRPADLPSPBALVDRKFWACRVGP
 LHSQDILGIVNADGELHSHNGAAGMQLCVDAQPLMLFSLGALTYVAILGSTI
 MTERGSPSLPCSPASTPTSPSALGIRLSDPSLSTGSGPLGSGAGCTAPNSVSPES
 PTPRLGQMSDECTCYEHAVDTVITYTGHMCLCYSCGLRLKKALHACCPICRRRIKD
 IIKYTRSS"

BASE COUNT 435 a 843 c 700 g 467 t
 ORIGIN

Query Match 23.9%; Score 399.6; DB 10; Length 2445;
 Best Local Similarity 56.2%; Pred. No. 2.8e-34;
 Matches 890; Conservative 0; Mismatches 634; Indels 60; Gaps 5;

99 GGTCTTGGGCGAGGCGCGCGCTTCCACGCGAGGCGCAAGAAAGTGCAGCTTGA 158
 171 GCTCCCGGCGACGCGCTCTTCCACCCCACTAAGGCTCCCAATCTCTATGGA 230
 159 CGGCACTGCGCGCGGCGCAACAGCGGCAAGCTTTCATATGCGTCACTTACGCA 218
 231 CTTCAAGCCACAAAGCGGTCAAGAGGCGAGGCTTCTCAATGCGCATCACTTCACTA 290
 219 GCGGCCAATCGCGGTGTACAGAGAGTGGCGCTGCGCTGTGCGCGCTGCGCTG 278
 291 CCGCCCGGTCTCATCTTACAGCAAGTCAAGGCTGAAGATCACCAAGAGCAATGCTGCTG 350
 279 GAGCGCGCGCTGCGCTTCCGCTTCCACGCGACGATCCGCTCATGAGCGCGCGCA 338
 351 GAGCGGCGCGCTGCGCTTCCGCTTCCACGCGAAGGACCTTCCCGCATCCACCGCGATC 410
 339 CATCCCAAGTACGCTGCGCGGACCTTGTCAAGCGCGCGCGGCTTATGCGCGCAAGCACT 398
 411 GCTGCCAAGTACGCTGCGCTTGTCAAGTGTCTCAAGTGTCTTGTGCGCAAAAGCAT 470
 399 GCCCGAAGACTGCGCGCTGCGCGACAGGCTGCTGCTACTGAGCGCGACCGCGACGCGCG 458
 471 GCTTGAAGAGTTTCCACAGAGGCGCAATCATGCTCTTGTGAGTGAACAAGAGGCGCG 530
 459 CGTGTCTACAGCGTGAAGAGCGGCGCGGCTCTTTCACCTGCGCGCGTGCCTGCGG 518

531 GATCTTCAACCGAGTCAATGAGTCAAGTCAATCTTCTTCAAGTGGAGTCCGACCGC 590
 519 CGGCGCGCTTGGGCGGCTCATTTAGTGTCAAGGATCAACGACAGAGTGCAGCTTCTGGA 578
 591 GACCGGCTTGGGCGGCTTGGAGAGCTTACGGCTCAACGGGGATGTCCAGCTGTCTAGA 650
 579 GAGCGCTTGGCTGACAGCTGACGCGCGCGGCTTCAAGCAGGCGCGCTTCAAGCGCTG 638
 651 CAGCGAGCTGG-----TGTGCGCGGACGTGCTGGGCGCGGCTCTTCAACGCGCT 701
 639 CTTGCGCGCGGACGACGACGAGCGGCGCAACTTGGACAAACAAGACTGCGAGAACACCA 698
 702 GCGCGCGCGTGGCTGCGGTGCGAGGCG-----GATGAGCGCG 740
 699 GATGATGACCAAGCTGAGGACACTGAGCGCTGAGCGCGCGCGCGCGCGCGCGCGCG 758
 741 CTTGTGATGAGCTTGTGCACTTCAAGTCCCGAGGCGCAAGCGCGCGCGCGCGCGCG 800
 759 CGCGCGCGCGCGCGCATTTCCGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 818
 801 GCTGCGCGCTGCGCGGATCCCGAGAACTCGCTCAATTCTAGGACAGCGCGCGCTGCC 860
 819 GCGGCTACTGAGGCGGACCTTGCCTTCAAGCAACAGCGCGCGCG---ACGTGAGCT 875
 861 GCGCGAGCTCGACGCGGACCTGCGCTTCAAGCGCTTGCAGCGCGCGCGCGCGCGCAT 920
 876 GTCCGCGCGACGCGCAAGTGGCTGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCT 935
 921 CTTGACAGAGAGACGCGGCGCGGCGGAGAGACGCGCGCGCGCGCGCGCGCTGCTCT 980
 936 CTTCCAGGCGCGCGCTGCGCGCGCGCGGAGAGCTTTCGTGAGTGGCGCGTCCGAGCT 995
 981 CACGAGCGCGCGGCTTGGCGGAGGACCATCTTCAAGGTCAAGCGCTGCGGAGTGG 1040
 996 GCGCGCGCGCGCGCGCTGCGCTTGCAGTCAAGCTGTCGACACCGCGCGCTGCTACGCGC 1055
 1041 CCGCGCGCGCGCGCGCTGTGTGTGCGGAGTCAACAGTGTGACCCCGGACGCTGCGCGC 1100
 1056 CAACAGAGTGCACCGCGCGACCGAGAGCGCTGTGACCGCAAGAGTACTGAGTGTGAGC 1115
 1101 GCGGAGCTTGCCTTCAAGCGCGCGGAGCGCTGTGTGACCGAGAGGTTCTGGGCGGTGG 1160
 1116 GCGCGCGCGCGCGCTGCGAGCGCGGAGCGCGCTCAAGCTTCAAGCTGCGCGCGCG 1175
 1161 TCGCGTGCAGGCGCTTGCACAGGCGGAGCATCTGAGGCTGTGTGCAACGCGAGCGG 1220
 1176 GAGGTGTCTTGTGAGTCAACAGGCGTCCGCGCGCGCGCGCTGTGTGTGACACGAC 1235
 1221 AGAGCTGACCTGAGTCAACAGGCGCGCGCGCGCATGCAAGCTGTGTGTGTGAGTGCCTC 1280
 1236 GCAAGCGCTGTGAGGCTTCTTCCGCGTGCAGCGCGCGCGCTGCGCGGCGAGCTGCTCTCT 1295
 1281 GAGGCGCTTGTGTGAGTCTTCTTCAAGCTGTGATGTGCGCATCAAGAGTCCGATCTCGG 1340
 1296 CGGTACCTTGCAGTCAAGCTTGCAGACCAAGCTTCAAGAGTCTTCAAGCG----- 1348
 1341 CTCGACATCATGATCAAGAGGAGTGCATCTCTCCCTGTGACCTGCTCTCACTCC 1400
 1349 -----GCTCCAGAGAGATGATGATGATTCATGATTCATGATTCATGATTCATGATTC 1397
 1401 AACCTACCCAGTGTGCGGTGATTCGCTCTCTGATCCCTGCTGACACCTGTGCGGTTC 1460
 1398 CTCTCGGATCTGATGATCTCTGTGTGATGAGCGCGCGCGCGCGCGCGCGCGCGCG 1457
 1461 TGGGCGCTTGTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1520
 1458 GTCCCGGT 1517
 1521 GTCAAGGAGT 1571
 1518 CTTGATGAGGAGT 1577

QY 1296 CGGTAACCTGAGTCCAGCCCTGCGACAGACTCCATCAGGTCCTCAGCG----- 1348
 DB 1559 CTCACCATCATGATGAGAAAGGGGGTGGCCCATCTCCCTGCTACCTGGCCCACTCC 1618
 QY 1349 -----GCTCCAGAGCATATGTATTCATATGATGACCTTGAATGTCAACAGTC 1397
 DB 1619 AACCTACACCAAGGCGCTGGGATCCGCTCTCTGACCCCTGCTCAGACCTGCGGTTTC 1678
 QY 1398 CTCCTGCGCATCTGATGATCCCTGGTGAAGGCCCCAGCTCCCGCTAGAGCCCCCGGT 1457
 DB 1679 TGGGCCCCTAGAGTGGCTCTGCTGAGAGGACAGCCCCCACTCCTGTAGGCTTCCCA 1738
 QY 1458 GTCCCCCGTCTTCCCAACCGAGCCGAGGACATCAAGATGGCGAGTGAACGCTGTG 1517
 DB 1739 GTACCCGGTGAACCCAGGCTGGGGCAGTG-----GAGTGAATGATGACCATTTGG 1789
 QY 1518 CTTTCATGGCGAGGTGAGACAGGTCATCTTACAGTGTGACACATGTGCTGTGCGACAG 1577
 DB 1790 CTATGAAACAGCAGTGAACAGTCATCTTACAGTGTGACACATGTGCTGTGCTACTTC 1849
 QY 1578 CTGCGGCTGCGGCTCAGAGGACCGGCGCTGCTGCGCCATCTGCGGCGGCGCAT 1637
 DB 1850 CTGTGCTGCGGCTCAGAGGACCGCTGACCGCTGCTGCGCCATCTGCGGCGGCGCAT 1909
 QY 1638 CAAGACGTCATTAGATCTTACAG 1661
 DB 1910 CAAGACATCATCAAGACCTTACCG 1933

RESULT 13

AF401228 3630 bp mRNA linear ROD 15-OCT-2001
 LOCUS Mus musculus neutralized (Neur1) mRNA, complete cds.
 DEFINITION AF401228
 ACCESSION AF401228
 VERSION AF401228.1 GI:15420882

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

JOURNAL

PUBMED

1. 3630
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1. 3630
 /gene="Neur1"
 156. 1829
 /note="similar to Drosophila melanogaster neutralized"
 /codon_start=1
 /product="neutralized"
 /protein_id="AAK97495.1"
 /db_xref="GI:15420883"
 /translation="MGQITRTYTHDSIGSPVPSHRCHNRKXKCPPTLSGGGUPAT
 PLTFPHTKGSOILMDLSHAKVROQSPFNATTFSPRPVLIIEQYRLKTKKQCCMSG
 ALRLEPTSDPSRIHDSLPKYACPDVLOSFGFMAKALPESBANNGNIIAPVDDKGR
 VPRINESAAMLPFGGRTVDPLMALVNYGLTRGQILDSLVLPDCLPSPSPALR
 RSLGCEADEARLSVSLCDLNPAGADGGAPAGCPITQNSLNSOHSALPAQIDGD
 LRFHAIKGAHVRIIDEGTVARLEHGRDERALVFTSRPVSVAETIFIKVTRSGGGRAG

BASE COUNT 667 A 1119 C 1082 G 762 T
 ORIGIN

Query Match 23.5%; Score 393.2; DB 10; Length 3630;
 Best Local Similarity 55.9%; Pred. No. 1.2e-33;
 Matches 886; Conservative 0; Mismatches 638; Indels 60; Gaps 5;

99 GGTCTGAGGAGGCGCGCGCTTCCAGCGAGGCGCAAGGCAAGAACTGCGGCTGGA 158
 DB 275 GCTCCCGGACAGCGCGCTGCTTCCAGCGGCACTAAAGGCTCCAGATCTCATGGA 334
 QY 159 CGGCACTCGGCGCGGCGCAACCGCGGCAAGAGCTTCTGCAATGGGTCAGCTTACGCA 218
 DB 335 CTTGAGCCACAAAGCGCGTCAAGAGGCAAGCGCTTCTGCAATGGGTCAGCTTACGCA 394
 QY 219 GCGGCGCAATCGGCTGATAGAGAGTGGCGCTGCGCTGAGTGGCGCTGCGCTGCTG 278
 DB 395 CGGCGGAGTGTCACTAGAGCAAGTCAAGCTGGAATGACCAAGAAAGCAATGCTGCTG 454
 QY 279 GAGCGGCGGCTGCGCTTCCAGCGGCAAGATCCGTCCTTACAGCGCCCAAGA 338
 DB 455 GAGCGGCGGCTGCGCTTCCAGCGGCAAGATCCGTCCTTCCAGCGCCCAAGTCC 514
 QY 339 CATGCCCAATAGAGCGGCGGAGCTGAGTCAAGCGGCGGCTTCTGAGGCGCAAGGCT 398
 DB 515 GCTGCGCAAGTACGCTGCGCTGAGCTGAGTGTCTGAGAGTGTCTTGGGCGCAAGGCT 574
 QY 339 GCGCGAAGACTGAGCGCTGAGCGGAGACCGGCTGAGCTTCTGAGGCGGCGCG 458
 DB 575 GCTGAGAGAGTGTGCAAGAGGAGCAATATGCTTCTGAGGAGTGTGCAAGAGGCGCG 634
 QY 459 GGTGTTCTACAGAGTGAACAGAGCGGAGCGGCTTCTTCACTGCGGCGTGGCG 518
 DB 635 GGTCTTCTACAGAGTGAACAGAGCGGAGCTGAGTGTCTTCTTCACTGAGGCGTGG 694
 QY 519 GCGGCGGCTGAGGCGCTCATTTAGTGTGAGAGGATGACCGAGAGGAGTGTGAGT 578
 DB 635 GAGCGGCTGAGGCGCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 754
 QY 579 GAGCGGCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 638
 DB 755 GAGCGGCTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAG 805
 QY 639 CTTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 698
 DB 806 GCGGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 844
 QY 699 GGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 758
 DB 845 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 904
 QY 759 GCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 818
 DB 905 GCTGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 964
 QY 819 GCGGCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875
 DB 965 GCGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1024
 QY 876 GTCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 935
 DB 1025 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1084
 QY 936 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 995
 DB 1085 CACGAGCGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1144
 QY 996 GCGGCGGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1055

Db	1145	GGGGGAGAGGGGGGGCGCTGCTCTTTCGGGGGTCAACACGTGTGACCTGGCAAGCTGGCGGC	1204
QY	1056	CAACGAGCTGCCCCCGCGACCAACGCGCTGCTGCACCGCAAGAGTACTGGGTGTGAC	1115
Db	1205	CGCGGACCTGCGCTTTCAGCCCCCGAGGCGCTGTGTGACCGCAAGGAGTTCTGGGCGGTGTG	1264
QY	1116	GCGGCGCGGGGCGCCGTGGCCGAGCGGGGGAACGCGTCAAGCTTCAACGTGCGGGCGCGGCG	1175
Db	1265	TGCGGTGCGCCGGGCGCTCTGCACAGCGGGGACATCTTGGGCGCTGTGTGTCAACGCGGACGG	1324
QY	1176	CGAGGTGCTCTGGGACATCAACGGGCGTCCGCGCGCGCGCGCTGTGTGTCTGCACACGAC	1235
Db	1325	AGAGCTGACCTGTGAGTCAACACGGGCGCGCGCGCGCATGACGTGTGTGTGATGTCTC	1384
QY	1236	GCAGCGGCTGTGGGCGCTTCTTTCGCGGTGCGCGCGCGCGGTGCGCGGACGTGCGTCTCT	1295
Db	1385	GCAGCGCGCTGTGAGTGTCTTTCAGCGCTCATGTGGCGCATCAACGAGGTCCGCATCTCGG	1444
QY	1236	CGGTACCTGCAATGCACGCTCGACCGCATCGACTTCATCAGAGTTCCTCAAGC-----	1348
Db	1445	CTCCACCATCATGACTGAACTGAACGGGGGTGGCCCATCTTCTCCGTGTCACTGTGCTCACTCC	1504
QY	1349	-----GCTCCGAGGACGAAATGATTCAGATATGACCTTCACTGTCAACAGTTC	1397
Db	1505	AACCTCACCCAGTGGCCCTGTGGGATATCCGCTCTTGAACCCCTGTCTCAGACATCTGGGTTTC	1564
QY	1398	CTCCTCGGCACTGTGAGTATCCCTGTGTGACGCGCCCCCAAGCTCCCGCTGAGCCCCCGGT	1457
Db	1565	TGGGCGCCCTTAGTGTGCTGTGCTGAGGGGACAGCCCCCACTCACTGTGAGCCTGTGCCGA	1624
QY	1458	GTCCCCCGTGTCTTCTCCACCGGAGCCGGCAGGACATCAAGAAATGGCGAGTGCAGGTGTG	1517
Db	1625	GTCAACCGGTGACCCCAAGGTCTGTGGCGCAATG-----GAGTGTGATTCACCAATTTG	1675
QY	1518	CTTGATGTGGCAGAGTGGACACGGGTATCTACGCTGTGAGACACATGTGCTGTGGACAG	1577
Db	1676	CTATGAACACGCAATGAAATACGATCATTTACAGGTGTGGCCACATGTGCTGTGTACTTC	1735
QY	1578	CTGCGGCGCTGTGGGCTCAAGCGCACAGGCCCGGCGCTGTGCCCATCTTGCAGGCGGCCAT	1637
Db	1736	CTGTGGCGCTGTGGCGCTCAAGAAAGGCGCTGTGCACGCGCTGTGCCCATCTTGTGCCCAT	1795
QY	1638	CAAGACGCTCATTTAAGATCTACAG	1661
Db	1796	CAAGACATCATCAAGACCTACCG	1819
RESULT 14			
LOCUS	E26375	1641 bp	DNA linear PAT 18-JUN-2001
DEFINITION	Neutralized protein, polynucleotide encoding said protein and		
ACCESSION	E26375		
VERSION	E26375.1	GI:13025071	
KEYWORDS	JP 199137257-A/2.		
SOURCE	unidentified		
ORGANISM	unidentified		
REFERENCE	1 (bases 1 to 1641)		
AUTHORS	Motomai, N., Hideo, N., Mitsuhiro, Y. and Hideyuki, S.		
TITLE	Neutralized protein, polynucleotide encoding said protein and		
JOURNAL	antibody recognizing said protein		
COMMENT	Patent: JP 199137257-A 2 25-MAY-1999;		
	SUMITOMO ELECTRIC IND LTD		
	OS Unidentified		
	PN JP 199137257-A/2		
	PD 25-MAY-1999		
	PF 14-NOV-1997 JP 1997313211		
	PR		
	PI MOTOMI NAKADA, HIDEO NAKAMURA, MITSUHIRO YOSHIDA, HIDEYUKI SAYA		
	PC C12N15/09, C07K14/47, C07K16/18, C12P21/02, C12Q1/68, G01N33/53//		
	PC C12N15/09, C12R1.91), (C12P21/02, C12R1.19), C12N15/00, (C12N15/00, PC		

FEATURES		BASE COUNT		ORIGIN	
source	location/Qualifiers	260	a	594	c
CC	Strandedness: Double;				
CC	Topology: Linear;				
PH	Key				
FT	source				
FT	1..1641				
FT	location/Qualifiers				
	1..1641				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
	260 a 594 c 498 g 288 t 1 others				
Query Match	23.4%; Score 392.2; DB 6; Length 1641;				
Best Local Similarity	55.9%; Pred. No. 2e-33;				
Matches	885; Conservative 0; Mismatches 639; Indels 60; Gaps 5				
99	GATCTGGGGGAGGGGCGCGGCTTCCAGCGCAGGGCCAAAGGCGAAGAACGTGCGGCTGA	158			
87	GCTCCGGGCGACGGCGCTGCTCTTCCACCCCGACATAGGGCTCCAGATCTCATGGA	146			
159	CGGCGACTCGGCGCGGCGCGACAGCGCGCAAGCTTCTGCAATGCGTCACTTCACGCA	218			
147	CCTGACCGCAAGGCGCGCTCAAGGCGAGCGCAGCTTCTGCAATGCGTCACTTCAGTAA	206			
219	GCGGCCCATCCGCGCTGTAAGACAGAGTGGCGCTGCGCTGTGCGCGCTGCGCTG	278			
207	CCGCGCGGTGCTCATCTACGAGCAAGTCAGGCTGTAAGTCAACAAAGAACATGCTGCT	266			
279	GAGCGGCGCGCTGCGCTTCGCTTCAACGCGCGACATTCCTGCTCATAGGCGCCAGGA	338			
267	GAGCGGGGCGCTGCGCATTTGCTTCAACGAGCAAGGACCTTCCGCGATCCACCGGACTC	326			
339	CATCCCCAAGTACGCGCTCCGCGACCTGATCAACGCGCGCGGCTACTGGGCGAAGGACT	398			
327	GCTGCGCCAGTAACGCTGCTCTTCAACGCTGATCTCAAGATGAGCTTCTGGGCGAAGACT	386			
399	GCCCGAGAACCTGGCGCTGCGCGACACAGGTGCTGCTCACTGGGCGGACCGCCAGCGCG	458			
387	GCTGAGGAGTTTGGCGAAGAGGCGAATCATCTTCCTTCTGCGGTGAGCAAGAGGCGCG	446			
459	GCTGTTCTACAGCGTGAACGAGCGGAGACCGGTGCTTTCCACTGCGCGGTGCGCGCG	518			
447	GCTCTTCAACCGGATCAATGAGTCAGCTGATATGCTTTCTTCAAGTGGGCTCCGACCGGT	506			
519	CGGCGCGCTCGGCGGCTCATTTGATGTCTAACGCGCATCAACGAGAGTGCAGCTTCTGGA	578			
507	GGAACCGGCTCTGGGCGCTTGTGAGAGCTTACCGGCTTCAACGCGGGGTGTCAGCTGTA	566			
579	GAGCGCTTCTGCTGACAGCTGACCGCGCGCGCGCTTCAAGCGCGCGCTTCAAGCGCTG	638			
567	CAGCGAGGCTGG-----TGTGCGCGCGACGTGCGCTGCGCGCGCTCTTCAACCGCGCT	617			
639	CCTGCGCGCCAGAGCGCAACGAGCGCGCGCACTTTCGACAAACGAGCTGAGAACACCA	698			
618	GCGGCGGCGCTGCTGCTGCGAGGAGCG-----GATGAAGCGCG	656			
699	GATGATGCGCAAGCTGCGGCGCACTGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCG	758			
657	CCTGCGGTGAGCGCTGTGCGACTCAACGTCGCGGAGCGCAAGCGGAGCAAGCGCGCG	716			
759	CGCGCGCGCGCGCGCATTCGCTGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	818			
717	GCTGCG	776			
819	GAGCGTACTGAGAGCGCGACTTCTTCAACGCAACGCGCGCGCGCGCGCGCGCGCGCG	875			
777	GAGCGAGCTGACCGCGCGCGCGCGCTTCAACGCGCGCGCGCGCGCGCGCGCGCGCG	836			
876	GTCG	935			
837	CCTGAGCAGGAGAGCGGTGCGCGCTGAGAGCAGCGCGCGCGCGCGCGCGCGCGCGCT	896			

Qy	936	TTCCGAGGCGCCGCTGCGCGGCGCCCGGAGAGCGCTTTGGTGGAGGGTGGGCGCTCCGCGGCT	995
Db	897	CACCAAGCCGCGCTTGAGCGTGGGCCGAGAACATCTTCAATCAAGGTCAAGCGCTCGGCGCG	956
Qy	996	GCGCGCGCCCGCGCGCGCTGGGCTTTCCGCGATCAAGTCGTGCGACCCCGGCGCGTTCACGCGC	1055
Db	957	GGGCGAGCGGGCGGGCGGTCTCTTTCGGGGGTCAACAGCGTGTAACTCCGTGGCAAGCTGGCGC	1016
Qy	1056	CAACGAGCTGCCCCCGCCAGACGCGCTGCTGACCTGGAAAGTACTGGGTGGTGGC	1115
Db	1017	GCGGAGCTGGCCCTTCACGCCCCCGAGGCCCTGATGGACCGCAAGGAGTTCTGGGCGGTGTG	1076
Qy	1116	GCGGCGCGGGCGCCGTCGCGAGCGGGCGGAGAGCGCTGACGCTTCAACGCTCGGGCGCGCGG	1175
Db	1077	TGCGGTGCGCCGGGCGCTCTGCAADAGGGGGAGACATCTGGGCGCTGGTGGTCAACGCGAGCG	1136
Qy	1176	CGACGTGCTCTTGAGCATCAACGCGCGTCCGCGCGCGCGCGCTGCTGTGCGTGAACACCA	1235
Db	1137	AGAGCTGACCTTGAGTCAACACGCGCGCGCGCGCGCATGACAGCTGTGCGTGAATGCTTC	1196
Qy	1236	GCAAGCGCTCTGGGCGCTTCTTTCGCGTGGCGGGCGGGCGGTCCGCGGCGACGTGCGTCTCT	1295
Db	1197	GCAAGCGCTCTGGAGTGCCTCTTCAGGCTGTCAGTGGCGCATCAAGCAGGTCCGCAATCTTCG	1256
Qy	1296	CGGTAACTTCGAGTCCAGCGCGTGGGAGCAACGACTTCATCAGGGTCCCTCAGCG-----	1348
Db	1257	CTCCACATCACTGACTGAACGGGGTGGCCCATCTTCTCCCTGCTACCTGCGCTTCACTCC	1316
Qy	1349	-----GCTCCAGAGACGATGATGATTCAGATATGACCTTCAGTGTCAACCACTC	1397
Db	1317	AACCTCAACCGATGCGCTGGGGCATTCGCGCTCTGAGACCCCGTGCACACACTGGCGTTC	1376
Qy	1398	CTCCTCGGCATCTGAATCATCCCTGGTGAAGCGCCCCCAAGCTTCCCGCTGAAGCCCCCGGT	1457
Db	1377	TGGGCGCCCTGAAGTGGCTCTGCTGAGGAGACAGCCCCCACTCACTGTGAGACTGGCCGA	1436
Qy	1458	GTCGCCCGGTGTTCTCCCAACGGGAGCGGAGAGGATCAAGATGGCGAGTGCACGAGTGG	1517
Db	1437	GCCACCGGTGACCCCGAGTCTGGGGCAG-----TGAAGTAGATGACATTTG	1487
Qy	1518	CTTCGATGGCAGGTGGAACAGGTCATCTACAGCTGTGAGACATGTGCTGTGCAAG	1577
Db	1488	CTATGAAACCGAAGTGAATCAAGTATCTTACAGCTGTGGCCACATGTGCGCTGTGTATTC	1547
Qy	1578	CTGCGGCTGCGGCTCAAGCAAGGCGCGGCGCTGTGCCCCATCTGCGGCGGCGCAT	1637
Db	1548	CTGTGGCTTGGCGCTCAAGAAAGGCGCTCAGCGCTGTGCCCCCATCTGCGCGCCCAT	1607
Qy	1638	CAAGGACGTACTTAAGATCTACAG	1661
Db	1608	CAGGACATCAACAGACTTACCG	1631

RESULT 15	AC110151/c	LOCUS	DEFINITION	AC110151	27731 bp	DNA	linear	HTG 22-SEP-2002
ACCESSION	AC110151							
VERSION	AC110151.4							
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.							
SOURCE	Rattus norvegicus (Norway rat)							
ORGANISM	Rattus norvegicus							
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.							
REFERENCE	1 (bases 1 to 27731)							
AUTHORS	Mueny, D., Merlie, H., Metzker, M., Lee, S., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Aryalalbechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, K., Barnstead, M., Behnemed, F., Bielawski, K., Blat, J., Blankenburg, K., Blyth, P., Brown, M.,							

Birtane, N., Buhay, C., Burch, P., Burrill, K., Cadenor, E.,
 Cavaden, V., Carter, K., Cavazos, I., Cesari, H., Centor, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M.L., Davis, C., Davy-Carrill, L., De Anda, C., Dedertich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinth, H., Ditya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Dublin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Fells, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Frazer, C.M., Gabriel, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, N.,
 Gunatacay, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, R., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpeth, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kows, C., Kratt, C.L., Lebow, H., Levan, T., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshuber, L., Louisaged, H., Lozada, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M. P., McNell, T. Z., Meenan, E.,
 Milosavljevic, A., Miner, G., Mnjia, B., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
 Nankya, C., Neal, D., Newton, S., Nguyen, N., Norris, S.,
 Nwaokemele, O., Okwunonu, G., Olarunpasogun, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,-L.,
 Puato, M., Quintero, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R.,
 Reilly, B., Reilly, M., Ren, X., Reuter, M., Richard, S., Riggs, F.,
 Rivers, C., Rodkey, T., Rojase, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shadratsbeym, A., Sisson, I., Sittler, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J.,
 Steinle, M., Strong, R., Sutton, A., Swartz, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tinjey, A., Trejce, Z., Uemami, K.,
 Valasek, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
 Williams, G., Willson, R., Wleciwyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhuesern, A., Weise, R., Smith, D. R., Holt, R.A., Smith, H.O.,
 Weinstein, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 277331)
 Morley, K.C.
 Direct Submission
 Submitted (10-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 277331)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (12-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Sep 19, 2002 this sequence version replaced g1:21743237.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu>
Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GRPU

[illegible]

9	409.2	24.4	2207	20	AAK77135	Human nerve mutant1
10	392.2	23.4	1641	20	AAK77136	Mouse nerve mutant1
11	391.6	23.4	1698	25	ABK75852	Mouse Neu1 cDNA #2
12	391.6	22.4	1749	25	ABK75851	Mouse Neu1 cDNA #1
13	380.8	22.7	1725	25	ABK75885	Rat Neu1 cDNA #1
14	380.6	22.7	1016	24	ABO43390	Oligonucleotide fcc
15	380.6	22.7	1016	24	ABO43391	Oligonucleotide fcc
16	355.6	21.2	642	24	ABO49468	Oligonucleotide fcc
17	355.6	21.2	642	24	ABO49482	Oligonucleotide fcc
18	324.8	19.4	642	24	ABO49482	Oligonucleotide fcc
19	324.8	19.4	642	24	ABO49483	Oligonucleotide fcc
20	264	15.8	1161	25	ABK75850	Human Neu1 cDNA #3
21	182.4	10.9	876	25	ABK75854	Mouse Neu1 cDNA #4
22	181.8	10.9	7905	22	ABK15199	Human nervous eye2
23	177	10.6	888	25	ABK75857	Rat Neu1 cDNA #3
24	176.8	10.6	1035	25	ABK75856	Rat Neu1 cDNA #2
25	176.4	10.5	1035	25	ABK75853	Mouse Neu1 cDNA #3
26	166.4	9.9	789	25	ABK75862	Human Neu3 cDNA
27	141.2	8.4	114955	20	AAK53491	Human adenosine A1
28	135.8	8.1	765	25	ABK75863	Mouse Neu3 cDNA
29	127	7.6	3974	23	ABL21564	Drosophila melanog
30	127	7.6	15366	23	ABL21564	Drosophila melanog
31	123.8	7.4	577	20	AAK77145	MO9925827 Seg ID N
32	122.4	7.3	949	23	AAK55348	DNA encoding novel
33	114.6	6.8	114955	20	AAK53491	Human adenosine A1
34	93	5.6	2307	24	ABK32842	Human encoding human
35	92.4	5.5	109519	22	AAK508692	Micromonospora DNA
36	91.2	5.4	441529	22	AAI199682	Mycobacterium tuberc
37	91	5.4	631	24	ABO43300	Oligonucleotide fcc
38	91	5.4	631	24	ABO43301	Oligonucleotide fcc
39	89	5.3	125401	22	AAAD17186	Streptomyces nourse
40	88.8	5.3	24379	18	AAAT91095	Streptomyces firenzi
41	88.8	5.3	24379	19	AAAT28925	Streptomyces roseo
42	83.4	5.0	3957	22	AAAO69686	HSV-2 immediate ear
43	83.4	5.0	154746	24	AAAD25519	Human herpesvirus
44	82.4	5.0	154746	24	AAAD25519	Human herpesvirus
45	82.4	4.9	493	25	ABK74792	Human cDNA sequenc

ALIGNMENTS

ID	ABX75858	standard;	CDNA, 1675 BP
AC	ABX75858;		
XX	31-MAR-2003	(first entry)	
DT			
XX	Human Neu2 cDNA, #1.		
DE			
XX	Human; mouse; rat; gene; ss; neuro		
KW	neurallised homology repeat; C3HC;		
KW	neurogenic function; transcript; calc		
KW	calcium-signal transducer; synap		
KW	tumorigenesis; organ developmen		
KW	neurogenic gene; nuclear transpo		
KW	inductive tissue interaction; cyto		
KW	neurodegenerative disease; cytos		
KW	tranquilliser; neuroprotective; Ne		
XX			
XX			
OS	Homio. sapiens.		
XX			
XX	US2002132233-A1.		
PN			
XX	19-SEP-2002.		
PD			
XX	14-MAR-2001; 2001US-0806387.		
XX			
PF			
PR	14-MAR-2001; 2001US-0806387.		
XX			
PA	(PALM/) PALM K.		

PA (TIMM/) TIMMUSK T.
XX
XX Palm K, Timmusk T;
XX
DR WP1: 2003-174081/17.
DR P-PSDB; ABUS6251.
XX
PT Novel neutralized polypeptide, New useful for controlling cell
PT proliferation and calcium signalling induced transcriptional processes
PT and treating depression, pain, anxiety, cancer and neurodegenerative
PT diseases -
XX
XX
XX Claim 6; Page 38-39; 75pp; English.
XX
XX The invention relates to a purified neutralised (Neu) polypeptide
CC comprising at least one neutralised homology repeat (NHR) domain and a
CC C3H4 RING-zinc finger domain, and a polynucleotide encoding the
CC polypeptide. The polynucleotide is useful for constructing a transformed
CC host cell that expresses a Neu protein. Neu has neurogenic function and
CC functions as a transcriptional regulator, as a calcium-signal transducer,
CC in cell signalling and synaptogenesis, in memory learning, RNA
CC processing, early immediate responses, death of specific cell
CC populations, nuclear hormone receptor signalling and axonal path-finding.
CC The Neu family of proteins presents a set of diagnostic and therapeutic
CC targets, especially for treating cancer. Manipulating Neu expression and
CC function is useful in controlling a variety of diseases including
CC depression, pain, anxiety, neurodegenerative diseases and cancer.
CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
CC invention.
CC
XX
SQ Sequence 1675 BP, 231 A, 648 C, 561 G, 235 T, 0 other;
Query Match 100.0%; Score 1675; DB 25; Length 1675;
Best Local Similarity 100.0%; Pred. No. 1.5e-257;
Matches 1675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGCAACACGGTGCACCGGACCCCTGCGACGCCGAGCCCAACCGGCGGCTCTGAGCC 60
DB 1 ATGGGCAACACGGTGCACCGGACCCCTGCGACGCCGAGCCCAACCGGCGGCTCTGAGCC 60
QY 61 ACCCGGCGTGTGCGGCGCCCGCGCCGAGCGACGCGCGGTCTGTGGGCGAGCGCGCGCC 120
DB 61 ACCCGGCGTGTGCGGCGCCCGCGCCGAGCGACGCGCGGTCTGTGGGCGAGCGCGCGCC 120
QY 121 TTCGACGCGGCAAGGCAAGGCAAGGCGGCGGTGAGAGGCGCACTGCGCGGCGGACACA 180
DB 121 TTCGACGCGGCAAGGCAAGGCAAGGCGGCGGTGAGAGGCGCACTGCGCGGCGGACACA 180
QY 181 CGGCGCAACAGCTTCTGCAATGAGCGTCAAGTTCAACGACGCGGCCATCCGCGTGTACGAG 240
DB 181 CGGCGCAACAGCTTCTGCAATGAGCGTCAAGTTCAACGACGCGGCCATCCGCGTGTACGAG 240
QY 241 CAGGTGCGGCTGCGCTGTGAGCGGTGCGCGCTTGTGAGAGCGGCGCTTGTGCGC 300
DB 241 CAGGTGCGGCTGCGCTGTGAGCGGTGCGCGCTTGTGAGAGCGGCGCTTGTGCGC 300
QY 301 TTCACGCGGCAAGATCGTGTGATGAGAGGCGGCAAGATCCCAAGTACGCTGCGCG 360
DB 301 TTCACGCGGCAAGATCGTGTGATGAGAGGCGGCAAGATCCCAAGTACGCTGCGCG 360
QY 361 GACCTGTGTCACGCGCGGCGCTAATGAGGCAAGGCACTGCGCGGAGAACTTGTGCGCGC 420
DB 361 GACCTGTGTCACGCGCGGCGCTAATGAGGCAAGGCACTGCGCGGAGAACTTGTGCGCGC 420
QY 421 GACAGCGGTGCTGCTTACTGTGGGCGGACCGGCAAGGCGCGGTGTTCTACAGCGTGAACGAC 480
DB 421 GACAGCGGTGCTGCTTACTGTGGGCGGACCGGCAAGGCGCGGTGTTCTACAGCGTGAACGAC 480

QY 481 GCGAGACCGGTGCTTCTCACTGCGGCGGTGAGCGGCGGCGGCTGTGAGCGCTCAAT 540
DB 481 GCGAGACCGGTGCTTCTCACTGCGGCGGTGAGCGGCGGCGGCTGTGAGCGCTCAAT 540
QY 541 GATGTCTAGGCAATCAACGAGAGGTGAGCTTCTGAGAGACGCTTGTGCTGACACGCTG 600
DB 541 GATGTCTAGGCAATCAACGAGAGGTGAGCTTCTGAGAGACGCTTGTGCTGACACGCTG 600
QY 601 ACGCGCGGCGCTGAGCGGCGGCTTCAAGCGGCTGAGCGGCTGAGCGGCGGCGGCGGCG 660
DB 601 ACGCGCGGCGCTGAGCGGCGGCTTCAAGCGGCTGAGCGGCTGAGCGGCGGCGGCGGCG 660
QY 661 GCGGCAACCTTCAACCAACGAGGTGAGGCAACGAGGTGTGAGGCAAGGTGAGGCGAC 720
DB 661 GCGGCAACCTTCAACCAACGAGGTGAGGCAACGAGGTGTGAGGCAAGGTGAGGCGAC 720
QY 721 CTGCGGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
DB 721 CTGCGGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 781 TCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
DB 781 TCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 841 CGCTTCAACGCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 841 CGCTTCAACGCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
QY 901 GCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
DB 901 GCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY 961 GAGAGCGCTTCTGAGGAGGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 961 GAGAGCGCTTCTGAGGAGGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
QY 1021 GGCATCAAGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 1021 GGCATCAAGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
QY 1081 GCGCTGCTGACCGGCAAGATGATCTGAGGTGTGAGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1081 GCGCTGCTGACCGGCAAGATGATCTGAGGTGTGAGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 1141 GCGGACGCGGCTCAGCTTCAAGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1141 GCGGACGCGGCTCAGCTTCAAGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
QY 1201 CGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
DB 1201 CGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
QY 1261 GTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
DB 1261 GTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
QY 1321 ACCAGGATCTCATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
DB 1321 ACCAGGATCTCATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
QY 1381 TTCAAGTGTCAACAGTCTCTCTGAGCATCTGAGTCAATCCCTGTGATACGCGCGCCAGCTCC 1440
DB 1381 TTCAAGTGTCAACAGTCTCTCTGAGCATCTGAGTCAATCCCTGTGATACGCGCGCCAGCTCC 1440
QY 1441 CCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
DB 1441 CCGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
QY 1501 GCGGAGTGAACGAGTGTGCTTCAATGAGGAGGTGAGCAAGGTATCTACAGTGTGAGCAAC 1560
DB 1501 GCGGAGTGAACGAGTGTGCTTCAATGAGGAGGTGAGCAAGGTATCTACAGTGTGAGCAAC 1560
QY 1561 ATGTGCTGTGCAACAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620

Db	934	GAGAGCGCTGTCGTTGGAAGTGAGGAGCCCGGAGGCTGGAGCGGCGCCGACGTGTGAGCTTCC	993
Qy	1021	GGCATTCACGTCGTGTGACACCCCGGAGTGTTACGCGCCCAACAGAGCTGTCCCGCCGACCCAGAAC	1080
Db	994	GGCATTCACGTCGTCTCGATTCCTGTGGCGCGCTGTGGCGGCATTCGAGCTGTCCCGCGAATCCGCT	1053
Qy	1081	GGCGTGTCTGCAGCCGCAAAAGAGTACTGGGTGGTGGCGCGCGCGGAGCCGTCGTCCGACGAGCGAGC	1140
Db	1054	GGCTGTCTGCAGACCGCAGAGGAATACTGGGTGGTGGCGCGCGCGCGCCGTCGTCCGACGAGC	1113
Qy	1141	GGCGACCGGCTCAGCTTCAGCGCTGCGAGCCCGGCGGAGCGTGTCTTGTGGCATCAACGGG	1200
Db	1114	GGCGACCGCATCAGCTTCACCGCTGCGACCGGCGGCGGACGTCGTCTGTGGCGGTGAAACGGG	1173
Qy	1201	CGTCCGCGCGGCGCGCGCTGTGGGTGACACACACGAGGGGAGCTGTGGAGCCCTTCTTGGCGC	1266
Db	1174	CGCCCGCGGGAGACGCTTGTGTGTGGTGAACCTTGACGGGGAGCTGTGGAGCCCTTCTTGGCT	1233
Qy	1261	GTGCGCGGCGGCGGTGCGCGGGGACAGCTGCGTCTCTCGGTACCCCTGCACTGACGCGCTGCG	1320
Db	1234	GTGCGCGGTTGGTGTGGCGGGATCAGCTTGGTCTCTCTGGGACCGTATCACTCGGATCTTGAG	1293
Qy	1321	ACCAACGACTCCATCAGAGGTCTCCCTCAGCGGCTTCCAGAGACATATGATTTCAAGATATGAC	1380
Db	1294	GCCACAACTCCATCAGAGGTCTCTTCAGTGACTTCAAGATATGACAGGATTTGAGACATGACC	1353
Qy	1381	TTCAAGTGTCAACCAAGTCTCTCTCGGACATCTGAGTCACTCCCTGATGACGGGCGCCGACGCTCC	1440
Db	1354	TTCCGGGGTCAACCAAGTCTCTCTCATTACAGCATATCAAGATGTCTGTGTGACAGCCGCCAGCTCC	1413
Qy	1441	CCGCTGAGCCCCCGGTGTCTCCCGGTGTCTTCTCCCAACCGAGCCCGGACAGGCAATCAAGAT	1500
Db	1414	CCACTGAGTCCCCCAGTGTCTCCCGGCTTCTCTGTGACCCGAGCCCGCGGACAGCGAAT	1473
Qy	1501	GGCGAGTGTCAAGGTGTGCTTGATGTGGAGGAGTGAACAGGTCATATCACGCTGTGACAC	1566
Db	1474	GGAGGTGTCAAGGTGTGCTTGAACAGCGAGGTGACAGCGTCACTTACAGCTGTGTGACAC	1533
Qy	1561	ATGTGCTCTGTGCCACAGCTGCGGCTGTGCGGCTCAAGCGACAGGCGCGGAGCTGTGCCCC	1620
Db	1534	ATGTGCTCTGTGCCACAGCTGCGGCTGTGCGGCTCCGAAAGCAGAGCAGGGCTGTGTCTCT	1593
Qy	1621	ATGTGCTCGCGCGGCGCCCATCAAGAGCGTCAATTAAGATCTACAGGCCCTAG	1668
Db	1594	ATGTGCTCGGTGGCGCCATCAAGAGTGTATCAAGATCTATAGGCGCTAG	1641

RESULT 3

ABX75859

ABX75859 standard; cDNA, 1129 BP.

ABX75859;

31-MAR-2003 (first entry)

Human Neu2 cDNA #2.

Human Neu2 gene; ss; neutralised polypeptide; Neu; NHR; cancer; neutralised homology repeat; C3HC4 RING-zinc finger domain; myogenesis; neurogenic function; transcriptional regulator; cell signalling; pain; calcium-signal transducer; synaptogenesis; memory learning; anxiety; tumorigenesis; organ development; central nervous system; depression; neurogenic gene; nuclear transport; Parkin-like protein; RNA processing; inductive tissue interaction; cytokine signaling; axonal path-finding; neurodegenerative disease; cytoskeletal antidepressant; analgesic; tranquilizer; neuroprotective; Neu1; Neu2; Neu3; Neu4.

Homo sapiens.

US2002132293-A1.

19-SEP-2002.

	PF	14-MAR-2001; 2001US-0808387.	
XX	PR	14-MAR-2001; 2001US-0808387.	
XX	PA	(PALM/) PALM K.	
XX	PA	(TIMW/) TIMMWSK T.	
XX	P1	Palm K., Timmusk T,	
XX	DR	MPI: 2003-174061/17.	
XX	DR	P-PDB; ABU56252.	
PT	PT	Newel neutralized polypeptide, Neu useful for controlling cell proliferation and calcium signaling induced transcriptional processes and treating depression, pain, anxiety, cancer and neurodegenerative diseases -	
XX	PS	Claim 6, Page 41, 75pp, English.	
CC	CC	The invention relates to a purified neutralised (Neu) polypeptide comprising at least one neutralised homology repeat (NHR) domain and a C3HC4 RING-zinc finger domain, and a polynucleotide encoding the polypeptide. The polynucleotide is useful for constructing a transformed host cell that expresses a Neu protein. Neu has neurogenic function and functions as a transcriptional regulator, as a calcium-signal transducer, in cell signalling and synaptogenesis, in memory learning, tumorigenesis, myogenesis and development of other organ systems and is related to repair and regeneration after injury to the central nervous system. Neu interacts with neurogenic genes, proteins implicated in nuclear transport and with Parkin-like proteins. Neu has the potential to interfere with inductive tissue interactions, cytokine signalling, RNA processing, early immediate responses, death of specific cell populations, nuclear hormone receptor signalling and axonal path-finding. The Neu family of proteins presents a set of diagnostic and therapeutic targets, especially for treating cancer. Manipulating Neu expression and function is useful in controlling a variety of diseases including depression, pain, anxiety, neurodegenerative diseases and cancer. Sequences ABR75648-ABR75664 represent Neu polynucleotides of the invention.	
CC	CC		
CC	CC		
SQ	SQ	Sequence 1129 BP; 158 A; 440 G; 372 C; 159 T; 0 other;	
Query Match		65.9%; Score 1103.6; DB 25; Length 1129;	
Best Local Similarity		98.8%; Pred. No. 8.2e-167;	
Matches 1112;		Conservative 0; Mismatches 14; Indels 0; Gaps 0	
DG	DG	550 GGCAATCACCGACAGAGTGCACCTTCGTGGAGAAGCGCCTTGCTGAACAAGCGGCCGG 609 4 GGCAACACGGTGCACCAGAACCTTCGACAGAGAAGCGCTTCGTGAACACTGAACGCCGCGG 63	
Gy	Gy	610 CGCCTTAGCCAGAGCCCCGTTTACAGAGCCCTGACCCCACAGACAAGACGGGCCAAG 669 64 GCCTCATAGCCAGGCCCGCTTAAGAGCCTGCTGACCGCCACAGAGACAAGACGGGCCAAAC 123	
Db	Gy	670 TTGCACAACAACAAGACTCAAAGAACAACAAGGTGTGTGACCAAGCTGAGGCACACTGACGTG 729 124 TTCACAACAACAAGACTCGAABAACAACAAGGTGTGTGACCAAGCTGAGGCACACTGAGCGCTG 183	
Dg	Gy	730 AGCGCGCGCCCCGGGCCCAACCGCCAGACCGACCGCGCGCCGCGCCATTCCGTGCGGCGCC 789 184 AGCGCGCGCCCCGGGCCCAACCGCCAGACCGACCGCGCGCCGCGCCATTCCGTGCGGCGCC 243	
Dd	Gy	790 GTGAGAGCGCCGCGCGCGCGCGGTGCGCGCGGCGGCACTGAGAGAGCGCACTGCGCTTCCAG 849 244 CTGAGAGCGCCGCGCGCGCGCGGTGCGCGCGGCGGCACTGAGAGAGCGCACTGCGCTTCCAG 303	
OY	Gy	850 GAACAACGCGGGCCCGAAGTGAAGCTGTGCGCCAAGCGCAAAAGTGAAGCTGTGCGCACCGCG 909 304 GAACAACGCGGGCCCGAAGTGAAGCTGTGCGCCAAGCGCAAAAGTGAAGCTGTGCGCACCGCG 363	
Ob	Gy	910 CCCAGAGCGCGCGCCCAAGCTGTCTTCCAGAGCGCCGCGTGTGCGCGCCGCGAGAGCGCTC 969 364 CCCAGAGCGCGCGCCCAAGCTGTCTTCCAGAGCGCCGCGTGTGCGCGCCGCGAGAGAGCTTC 423	

QY 970 TTCTGTGAGAGTGGGCGCTCCGGGGGCTGGCGGGCGCGCGGCTTGGGATCAAG 1029
 DB 424 TTCTGTGAGAGTGGGCGCTCCGGGGGCTGGCGGGCGCGCGGCTTGGGATCAAG 483
 QY 1030 TCGGTGAGAGTGGGCGCTCCGGGGGCTGGCGGGCGCGCGGCTTGGGATCAAG 1089
 DB 484 TCGGTGAGAGTGGGCGCTCCGGGGGCTGGCGGGCGCGCGGCTTGGGATCAAG 543
 QY 1090 GACCGCAAGAGTACTGGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1149
 DB 544 GACCGCAAGAGTACTGGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 603
 QY 1150 CTGAGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1209
 DB 604 CTGAGCTTCAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 663
 QY 1210 GGGCGCGCTGTGTGGGTGACACCAAGCGAGCGCTTGGGCTTCTTGGCGTGGCG 1269
 DB 664 GGGCGCGCTGTGTGGGTGACACCAAGCGAGCGCTTGGGCTTCTTGGCGTGGCG 723
 QY 1270 GGGCGTGGGCGCGAGCTGCTCTCTCGGTACCTTGCAGTCCAGCGCTTGGCGAG 1329
 DB 724 GGGCGTGGGCGCGAGCTGCTCTCTCGGTACCTTGCAGTCCAGCGCTTGGCGAG 783
 QY 1330 CCATCAGAGTCTCTCAGCGCTCCCAAGAGCAATGATGATTTGATATGATCTTCA 1389
 DB 784 CCATCAGAGTCTCTCAGCGCTCCCAAGAGCAATGATGATTTGATATGATCTTCA 843
 QY 1390 AACAGTCTCTCGGAGCTGAGTATCTCGGTGAGCGCGCGCGCGCGCGCGCGCG 1449
 DB 844 AACAGTCTCTCGGAGCTGAGTATCTCGGTGAGCGCGCGCGCGCGCGCGCGCG 903
 QY 1450 CCCCCGCTGTCCTCCGCTTCTCCCAAGCGAGCGCGCGCGCGCGCGCGCGCGCG 1509
 DB 904 CCCCCGCTGTCCTCCGCTTCTCCCAAGCGAGCGCGCGCGCGCGCGCGCGCGCG 963
 QY 1510 ACGGTGTCTTGTGATGCGAGAGTGAACAGGCTCATCTTCAAGTGTGACATGTGCTG 1569
 DB 964 ACGGTGTCTTGTGATGCGAGAGTGAACAGGCTCATCTTCAAGTGTGACATGTGCTG 1023
 QY 1570 TGCGCAGCTGCGGCGCTGCGGCTCAAGGAGCAAGCGCGCGCGCGCGCGCGCG 1629
 DB 1024 TGCGCAGCTGCGGCGCTGCGGCTCAAGGAGCAAGCGCGCGCGCGCGCGCGCG 1083
 QY 1630 CGGCGCATCAAGAGCTGATTAAGATCTACAGGCGCATGCTTACGCG 1675
 DB 1084 CGGCGCATCAAGAGCTGATTAAGATCTACAGGCGCATGCTTACGCG 1129
 RESULT 4
 ID ABX75860 standard; cDNA; 955 BP.
 AC ABX75860;
 XX
 DT 31-MAR-2003 (first entry)
 XX
 DE Human Neu2 cDNA #3.
 XX
 KW Human; mouse; rat; gene; 89; neutralised polypeptide; Neu; NHR; cancer;
 KW neutralised homology repeat; C3HC4 RING-zinc finger domain; myogenesis;
 KW neurogenic function; transcriptional regulator; cell signaling; pain;
 KW calcium-signal transducer; synaptogenesis; memory learning; anxiety;
 KW tumorigenesis; organ development; central nervous system; depression;
 KW neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;
 KW inductive tissue interaction; cytokine signaling; axonal path-finding;
 KW neurodegenerative disease; cytoskeletal; antidepressant; analgesic;
 KW transduces; neuroprotective; Neu1; Neu2; Neu3; Neu4.
 KW
 OS Homo sapiens.
 XX
 XX
 PN US2002132293-A1.

XX
 PD 19-SEP-2002.
 XX
 PP 14-MAR-2001; 2001US-0808387.
 XX
 PR 14-MAR-2001; 2001US-0808387.
 XX
 PA (PALM/) PALM K.
 XX (TMM/) TMMUSK T.
 XX
 PI Palm K, Tilmusk T;
 XX
 XX WPI; 2003-174091/17.
 DR P-PSDB; ABUS6253.
 DR
 PT Novel neutralized polypeptide, Neu useful for controlling cell
 PT proliferation and calcium signaling induced transcriptional processes
 PT and treating depression, pain, anxiety, cancer and neurodegenerative
 PT diseases -
 XX
 PS Claim 6; Page 43; 75pp; English.
 XX
 XX The invention relates to a purified neutralised (Neu) polypeptide
 CC comprising at least one neutralised homology repeat (NHR) domain and a
 CC C3HC4 RING-zinc finger domain, and a polynucleotide encoding the
 CC polypeptide. The polynucleotide is useful for constructing a transformed
 CC host cell that expresses a Neu protein. Neu has neurogenic function and
 CC functions as a transcriptional regulator, as a calcium-signal transducer,
 CC in cell signaling and synaptogenesis, in memory learning,
 CC tumorigenesis, myogenesis and development of other organ systems and is
 CC related to repair and regeneration after injury to the central nervous
 CC system. Neu interacts with neurogenic genes, proteins implicated in
 CC nuclear transport and with parkin-like proteins. Neu has the potential to
 CC interfere with inductive tissue interactions, cytokine signaling, RNA
 CC processing, early immediate responses, death of specific cell
 CC populations, nuclear hormone receptor signaling and axonal path-finding.
 CC The Neu family of proteins presents a set of diagnostic and therapeutic
 CC targets, especially for treating cancer. Manipulating Neu expression and
 CC function is useful in controlling a variety of diseases including
 CC depression, pain, anxiety, neurodegenerative diseases and cancer.
 CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
 CC invention.
 XX
 XX Sequence 955 BP; 148 A; 355 C; 304 G; 148 T; 0 other;
 SQ
 Query Match 34.5%; Score 577.2; DB 25; Length 955;
 Best Local Similarity 99.5%; Pred. No. 3.1e-83;
 Matches 579; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATGGGCAACAGGTGACACCGGACCTTGCACAGCCCAAGCGCGCGCTCTTGGCC 60
 DB 1 ATGGGCAACAGGTGACACCGGACCTTGCACAGCCCAAGCGCGCGCTCTTGGCC 60
 QY 61 ACCGGCGGTGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 DB 61 ACCGGCGGTGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 QY 121 TTCCACGCGCAAGGCAAGAGCAAGAGCGTGGAGTGAACGCGCACTGCGCGCGGCA 180
 DB 121 TTCCACGCGCAAGGCAAGAGCAAGAGCGTGGAGTGAACGCGCACTGCGCGCGGCA 180
 QY 181 CCGGCGCAACAGCTTCTGCAATGCGGTCACTTCAACGAGCGCGCGCGCGCGCGCG 240
 DB 181 CCGGCGCAACAGCTTCTGCAATGCGGTCACTTCAACGAGCGCGCGCGCGCGCGCG 240
 QY 241 CAGGTGCGGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 DB 241 CAGGTGCGGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 QY 301 TTCCACGCGCAAGTCCGTGCTCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 DB 301 TTCCACGCGCAAGTCCGTGCTCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 360

Oy	361	GACCTGGTCACCGCGCCGCGGCTACTAGTGGCCCAAGGGAACGCGCCGGAACACTGGCGCTGGGC	420
Db	361	GACCTGGTCACCGCGCCGCGGCTACTAGTGGCCCAAGGGAACGCGCCGGAACACTGGCGCTGGGC	420
Oy	421	GACACGGGTGCTGGCGCTTACTAGTGGGCCGACCGGCCCAACCGCGCGTGTCTTACAGCGCTGAACGAC	480
Db	421	GACACGGGTGCTGGCGCTTACTAGTGGGCCGACCGGCCCAACCGCGCGTGTCTTACAGCGCTGAACGAC	480
Oy	481	GCGCAGCGCGTGCCTCTTCCACTGACGCGCGCGTGGCGCGCGCGCTCTGGGCGCTCATTT	540
Db	481	GCGCAGCGCGTGCCTCTTCCACTGACGCGCGCGTGGCGCGCGCGCTCTGGGCGCTCATTT	540
Oy	541	GATGCTTACGGGATCACCAGCAGAGGTGCAGCTTCTGGAGAGC	582
Db	541	GATGCTTACGGGATCACCAGCAGAGGTGCAGCTTCTGGAGTAC	582

RESULT 5
ABQ43992/c
ID ABQ43992 standard; DNA; 1016 BP.
XX AC
XX ABQ43992;
XX
XX 12-JUL-2002 (first entry)
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 30563.
XX DE
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX
XX 01-SEP-2001; 2001WO-EP10074.
XX PF
XX 01-SEP-2000; 2000DE-1043826.
XX PR 05-SEP-2000; 2000DE-1044543.
XX
XX (EPiG-) EPIGENOMICS AG.
XX PA
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX PI WPI; 2002-371829/40.
XX DR
XX
XX PT Determining the degree of cytosine methylation in genomic DNA, useful
XX PT for diagnosis and prognosis, comprises selective hybridization of
XX PT amplicons from chemically treated DNA -
XX
XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridized to two classes, each with at least one
XX member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
XX and the degree of hybridization to both classes is determined from the
XX label on the amplicon. From the ratio of labels hybridized to the two
XX classes of oligomers, the degree of methylation is calculated. The method
XX is used: (i) for diagnosis and/or prognosis of side effects of
XX therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
XX of the central nervous, cardiovascular, gastrointestinal and respiratory
XX systems etc., particularly by detecting mutations or single nucleotide
XX polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
XX types and for investigating cell differentiation. The method allows the
XX methylation status of many C residues to be determined simultaneously.
XX ABQ43992-ABQ54121 represent genomic DNA sequences used to illustrate the

CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 1016 BP, 132 A, 154 C, 421 G, 306 T, 3 other;

Query Match	25.8%	Score 432.4	DB 24	Length 1016
Best Local Similarity	78.0%	Pred. NC. 3e-60		
Matches 581, Conservative	0	Mismatches 159	Indels 5	Gaps 5

QY	578	AGAGCGCTTTCGTGACACGCTGACGCGCGCGGCTCAAGCAGGCGCTTCAAGCGCT	637
Db	912	AAAACGCGCTTCGTACAAACG-7AAACGCGCGCGCTTCAACAAACCGCTTCAAGCGCT	854
QY	638	GCCGCGCGCCAGGAGCCACGACGCGGCGCAACTTCGACAAACGAGCTCGAGAACAC	697
Db	853	ACCTACCGCGCCGACACACGACGCGACCAACTTCGACAAACGAGCTCGAGAACAC	794
QY	698	AGGTGTGTGCGCAGCTGAGGCGACCTGAGCGCTGAGCGCGCGCGCGCGCGCGCGCG	757
Db	793	AAATATATACCAAACTAAACCACTCG-ACATAACCGCGCGCGCGCGCGCGCGCGCG	736
QY	758	ACGCGCGCGCGCGCGCGCGCGCATTC-CCGTGCGCGCGCGCGCGCGCGCGCGCGCG	816
Db	735	ACGCGCGCGCGCGCGCGCGCGCATTCCTGTATAGAAACCCGTAAACCGCGCGCGCGCG	676
QY	817	CCGCGCGCTACTGAGAGCGCGACCTGCGCTTCACGACCAACACGCGGCGCGACGTGAGCTG	876
Db	675	CGAGCGCTACTTAAATAACCGACCTACGCTTCACCGCACACGCGGACCGGACGTAAACCTTA	616
QY	877	TCGGCGCGACCGCAAAATGTAGCTGTGCGCACCGCGGCGCGACCGCGCGCGCTGCTTC	936
Db	615	TCGACCGACCGCAAAATTAACCTTAACGACAAACGACCGGAAACGACGCGACTAATCTTC	556
QY	937	TCCG-AGCGCGCGCTGCGCGCGCGCGGAGAGCGCTTCGTGTGAGGTGTGCGCGCGCT	995
Db	555	TCCGAAATCCCCGCTACGACAAACCGCAAAACCTTCCTGTAAATAAATACCGTCCGAAATC	496
QY	996	GGCGCGCGCGCGCGCGCGCTGCGGCTTCGCGCATCAGCGTGTGCGGACCGCGGCGGTACGCGC	1055
Db	495	AACGACGCGCGCGCGCGCTTAACCTTCGACATCACTGTATGACACCGGAACTATCTAGAC	436
QY	1056	CAACGAGCTGCGCGCGCGACCGACGCGCGCTCTCGACCGCAAAAGTACTGTGTGTGCGC	1115
Db	435	CACGAACTACCGCGCGCGACCGAAACCGCGCTACTCGACCGCAAAATACTAAATTAAC	376
QY	1116	GCGCGCGCGGCGCGCGCGCGGAGCGCGCGCGCGCGCTCACTTCACTGCGCGCGCGCGG	1175
Db	375	GCGCGCGGAAACCGGTACCGAAACGACGACGCGCGCTCAACTTCACTGACACCGGACGGA	316
QY	1176	CGACGTGCTCTGTGAGCATCAACGCGCGCTTCGCGCGCGCGCGCGCTGTGTGTGTGACACAC	1233
Db	315	CGACGTACTCTTAAACATCAACGACGTCGCGCGCGCGCGCTTACTATAGTGTGACACAC	256
QY	1236	GCAAGCGCTGTGAGGCGCTTCCTTCGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCTCT	1295
Db	255	GCAACGCGCTTAACCTTCTGTGCGCGTACGCGGACGACGTGTGCGGAAACAACTACGTCTCT	196
QY	1296	CGGTACCGTGTGACGTGACGCGCGCTTCG	1320
Db	195	CGATTAATCCCGGACCGCGCGGTACG	171
RESULT 6			
ABQ43993			
ID ABQ43993 standard, DNA, 1016 BP.			
XX ABQ43993;			
XX			
DT 12-JUL-2002 (first entry)			
XX			
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30584.			
KW Human, cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;			

KM drug; side effect; cancer; central nervous system; cardiovascular;
 KM gastrointestinal; respiratory system; single nucleotide polymorphism;
 KM SNP; cell differentiation; ds.
 XX Homo sapiens.
 OS
 PN W0200218632-A2.
 PD 07-MAR-2002.
 PF 01-SEP-2001; 2001WO-BP10074.
 PR 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 PA (EPIC-) EPIGENOMICS AG.
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 DR WPI; 2002-371829/40.
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc.; particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ1310-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 SQ Sequence 1016 BP; 306 A; 421 C; 154 G; 132 T; 3 other;
 Query Match 25.8%; Score 432.4; DB 24; Length 1016;
 Best Local Similarity 78.0%; Pred. No. 3e-60;
 Matches 581; Conservative 0; Mismatches 159; Indels 5; Gaps 5;

QY 877 TCGCGGACCGCAAGTGGCTGCGACACGCGGACCGGCGGCGGCGGCGGCTTC 936
 DB 402 TCGACGACCGCAAAATTAACCTACGACMAAGACCGAACGACGACGCTAATCTTC 461
 QY 937 TCCG-AGCGCCCGCTGCGGCGCGGAGAGACCTCTTCGTGAGAGTGGCCGCGGAGCT 995
 DB 462 TCCGAANAGCCCGCTACGACGACGCGGAAACCTCTTGTAATAATAACCGTCGAAACT 521
 QY 996 GCGGCGCGCGCGCGCGCTTGGCTTCCGATCAGTGTGCGACCCGCGGCGCTGACGCC 1055
 DB 522 AACGACGCGCGGACCGGCTTACGACATCAAGTGTGACGACCGGAACTGACGAC 581
 QY 1056 CAACGAGCTGCGCGCGGACCGGACGCGGCTGACCGGAAAGAGTACTGGGTGGC 1115
 DB 582 CAACGAGCTGCGCGGACCGGACCGGCTGACCGGAAAGAGTACTGAAATATATATAC 641
 QY 1116 GCGCGCGCGGCGCGGCTGCGGACCGGCGGCGGCTGACGCTTCAAGTGGCGCGGCGG 1175
 DB 642 GCGCGCGGACCGGCTACCGGACGACGACGACGACGCTCACTTCAAGTGGCGGCGG 701
 QY 1176 GAGCTGCTCTTGGGATCAACGCGGCTCGCGGCGGCGGCTGCTGTGCTGACACAC 1235
 DB 702 CGAGTACTCTTAAACATCAACGACGTCGCGGACCGGCTACTATACGTCGACACAC 761
 QY 1236 GCGGCGCTGCGGCTTCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTCT 1295
 DB 762 GCAAGCGCTTAAACCTTCTTCCGCGGACGACGACGACGACGACCACTACGCTCT 821
 QY 1296 CGGTACCTGACGTCGACGCGCTGCG 1320
 DB 822 CGATTAATCCCGACCGCGGTACG 846
 RESULT 7
 ID ABX75849 standard; cDNA; 1674 BP.
 XX
 AC ABX75849;
 XX
 DT 31-MAR-2003 (first entry)
 XX
 DB Human Neu1 cDNA #2.
 XX
 KM Human; mouse; rat; gene; 88; neutralised polypeptide; Neu; NHR; cancer;
 KM neutralised homology repeat; C3Hc4 RING-zinc finger domain; myogenesis;
 KM neurogenic function; transcriptional regulator; cell signalling; pain;
 KM calcium-signal transducer; synaptogenesis; memory learning; anxiety;
 KM tumorigenesis; organ development; central nervous system; depression;
 KM neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;
 KM inductive tissue interaction; cytokine signalling; axonal path-finding;
 KM neurodegenerative disease; cytoskeletal; antidepressant; analgesic;
 KM tranquiliser; neuroprotective; Neu1; Neu2; Neu3; Neu4.
 XX
 OS Homo sapiens.
 XX
 PN US2002132293-A1.
 XX
 PD 19-SEP-2002.
 PF 14-MAR-2001; 2001US-0808387.
 PR 14-MAR-2001; 2001US-0808387.
 PA (PALM/) PALM K.
 PA (TIMM/) TIMMUSK T.
 XX
 PI Palm K, Timmusk T;
 XX
 DR WPI; 2003-174081/17.
 DR P-PSDB; ABUS6242.
 XX
 PT Novel neutralised polypeptide, Neu useful for controlling cell

PT proliferation and calcium signaling induced transcriptional processes
PT and treating depression, pain, anxiety, cancer and neurodegenerative
PT diseases -

PS Claim 6; Page 21-22; 75pp; English.

XX The invention relates to a purified neutralised (Neu) polypeptide
XX comprising at least one neutralised homology repeat (NHR) domain and a
CC C3H4 RING-zinc finger domain, and a polynucleotide encoding the
CC polypeptide. The polynucleotide is useful for constructing a transformed
CC host cell that expresses a Neu protein. Neu has neurogenic function and
CC functions as a transcriptional regulator, as a calcium-signal transducer,
CC in cell signalling and synaptogenesis, in memory learning,
CC tumorigenesis, myogenesis and development of other organ systems and is
CC related to repair and regeneration after injury to the central nervous
CC system. Neu interacts with neurogenic genes, proteins implicated in
CC nuclear transport and with Parkin-like proteins. Neu has the potential to
CC interfere with inductive tissue interactions, cytokine signalling, RNA
CC processing, early immediate responses, death of specific cell
CC populations, nuclear hormone receptor signalling and axonal path-finding.
CC The Neu family of proteins presents a set of diagnostic and therapeutic
CC targets, especially for treating cancer. Manipulating Neu expression and
CC function is useful in controlling a variety of diseases including
CC depression, pain, anxiety, neurodegenerative diseases and cancer.
CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
CC invention.

XX Sequence 1674 BP; 257 A; 634 C; 510 G; 273 T; 0 other;

Query Match 24.4%; Score 409.2; DB 25; Length 1674;
Best Local Similarity 56.6%; Pred. No. 1.4e-56;
Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

QY 99 GGTCTCGGCGAGGCGCCCGCTTCCACGCGGAGGCCAAGCAAGACGTGGGTGGA 158
DB 120 GCTCCAGGACACGCCCTGCTTCCACCGCACACCAAGGCTCCCAATCTCATGGA 179
QY 159 CGGCACTCGCGCGCGGACCAACGCGGCAACAGCTTCTGCAATAGGCTCATGTCAGCA 218
DB 180 CTTGAGCCACAAAGGCTGTCAAGAGGAGGCGCAAGCTTCTGCAAGCCATCTTACGAA 239
QY 219 GCGGCCATCCGCGCTGACGAGCAGGTGCGGCTGCGCTGAGGCGCGTCCGCTGCTG 278
DB 240 CGCGCGGCTCTCATCTAGAGCAGTCAAGCTGAAGATCACCAAGAGAGGCTGCTG 299
QY 279 GAGCGCGCGCTGCGCTTCCACCGGACGATCCGCTCATGAGCGCCAGCA 338
DB 300 GAGCGGCGCTGCGCTTCCACCGGACGATCCGCTCATGAGCGCGCTGCTG 359
QY 339 CATCCCAAGTACGCGTCCGAGCGGATCAAGCGGCGGCTGAGGCGCAAGGACAT 398
DB 360 GCTGCCAAGTACGCGTCCGAGCGGATCAAGCGGCGGCTGAGGCGCAAGGCGCT 419
QY 399 GCCCGAAGACTGCGCTGCGGACACGCTGCTGAGCTGAGGCGCGAGCGCG 458
DB 420 GCTGAGGAGTTGCAATGAGGAGCAATCATCGATCTGAGTGAACAAGAGGCG 479
QY 459 CGTGTCTACAGGTGAGCAAGCGGCGGCTGCTTCCATGCGGCTGAGCGCTG 518
DB 480 TGTCTTCAACCGGATCAAGCACTCGGCTGTTATGCTGTTCTTACGAGGCTCGCAAG 539
QY 519 CGGCGCGCTTGGGCGCTCATGATGTCTACGCGCATCAAGCAAGGTGAGCTTCTG 578
DB 540 GAGCGGCTTGGGCGCTTGGTGAAGCTTACGCGCTCAAGCGGCGGCTCAAGTGTG 599
QY 579 GAGCGGCTTGGTGAAGCTGAGCGCGCGGCTTCAAGCAAGCGCTTCAAGCGCTG 638
DB 600 TAGGAGCTGG-----TGTCTCCGAGCTGTGCGGCGCGCTTCAAGCGCTG 650
QY 639 CTTGCGGCGGAGCAAGCGCGGCGCACTTTCAGCAACAAGAGCTGAGAAACA 698
DB 651 GCGGCGGCGCTGCTGCGGCGGAGGCGAGCAAGCGCGCTCTCTGCTG----- 699

QY 699 GGTGAGCCAAAGTGGGCACTGGGCACTGGGCGCGCGCGCGCGCGCGCGCGCA 758
DB 700 -----AGCTTATGCACTCAAGTGGCGGCGCGGAGCGGAGCGCGCGCG 749
QY 759 GCGCGCGCGCGCGCATTTCCGTGCGGCGCGCGCGCGCGCGCGCGCGCGCTG 818
DB 750 GCG 809
QY 819 GCGCGTATGAGAGCGGAGCTGCGCTTCCAGCAACAAGCGGCGCGCGCGCG 875
DB 810 GCGCGAGCTTCAAGCGGAGCTGCGCTTCCAGCGCGCGCGCGCGCGCGCGCG 869
QY 876 GTGCGCGAGCGCAAGTGGCTGCGCAAGCGGCGCGCGCGCGCGCGCGCGCT 935
DB 870 CTTGAGAGAGAGAGCGGTGCGCGCGGTGAGAGCAAGCGCGCGCGCGCGCT 929
QY 936 CTTGAGAGAGAGAGCGGTGCGCGCGGTGAGAGCAAGCGCGCGCGCGCGCT 995
DB 930 CACCAAGCG 989
QY 996 GCG 1055
DB 990 GCG 1049
QY 1056 CAAGAGCTGCG 1115
DB 1050 GAGCGAGCTGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1109
QY 1116 GCG 1175
DB 1110 CGCGGTGCG 1169
QY 1176 CGAGGTGCTCTGAGGATCAAGCGGCGGTGCGCGCGCGCGCGCGCGCGCG 1235
DB 1170 CGAGGTGCACTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1229
QY 1236 GAGGCGCTTGGGCGCTTCTTCTGCGCGCGCGCGCGCGCGCGCGCGCG 1295
DB 1230 GAGGCGCTTGGGCGCTTCTTCTGCGCGCGCGCGCGCGCGCGCGCGCG 1289
QY 1296 CGGTACCTGAGGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1346
DB 1290 CTCGACTATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1349
QY 1347 -----CGGCTCCAGAGAGATGATGATGATGATGATGATGATGATGAT 1397
DB 1350 AACCTGCG 1409
QY 1398 CTCCTGCGGATCTGAGTCAATCCCTGAGTCAAGCGCGCGCGCGCGCGCG 1457
DB 1410 TGGCGCTTGGGAGCTGCTGCTGAGTCAAGCGCGCGCGCGCGCGCGCG 1469
QY 1458 GTCCCGCGGTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1517
DB 1470 GTTCGCAAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1520
QY 1518 CTTGAGTGGAGAGTGAAGCAAGGTATCTTCAAGTGAAGCAAGTGAAGTGA 1577
DB 1521 CTTGAGTGGAGAGTGAAGCAAGGTATCTTCAAGTGAAGCAAGTGAAGTGA 1580
QY 1578 GTGCGCGCTTGGGCTTCAAGGAGCAAGCGCGCGCGCGCGCGCGCGCGCG 1637
DB 1581 CTTGAGCGCGCGCTTCAAGGAGCTTGAAGCGCGCGCGCGCGCGCGCGCG 1640
QY 1638 CAAGAGCTCATTAAGATCTACAG 1661
DB 1641 CAAGAGCTCATTAAGATCTACAG 1664

RESULT 8
ABX75848
ID ABX75848 Standard; cDNA; 1725 BP.
XX

AC ABX75848;
 XX 31-MAR-2003 (first entry)
 XX Human Neu1 cDNA #1.
 DE
 XX
 KM Human; mouse; rat; gene; ss; neutralised polypeptide; Neu; NHR; cancer;
 KM neutralised homology repeat; C3HC4 RING-zinc finger domain; myogenesis;
 KM neurogenic function; transcriptional regulator; cell signalling; pain;
 KM Calcium-Signal transducer; synaptogenesis; memory learning; anxiety;
 KM tumorigenesis; organ development; central nervous system; depression;
 KM neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;
 KM inductive tissue interaction; cytokine signalling; axonal path-finding;
 KM neurodegenerative disease; cytostatic; antidepressant; analgesic;
 KM Tranquilliser; neuroprotective; Neu1; Neu2; Neu3; Neu4.
 XX
 XX Homo sapiens.
 OS
 XX US2002132293-A1.
 PN
 XX 19-SEP-2002.
 PD
 XX 14-MAR-2001; 2001US-0808387.
 PF
 XX 14-MAR-2001; 2001US-0808387.
 PR
 XX 14-MAR-2001; 2001US-0808387.
 PA
 XX (PALM)/ PALM K.
 PA (TIMM)/ TIMMUSK T.
 PI
 XX Palm K, Timmusk T;
 DR WPI; 2003-174081/17.
 DR P-PSDB; AB056241.
 XX
 PT Novel neutralized polypeptide, Neu useful for controlling cell
 PT proliferation and calcium signaling induced transcriptional processes
 PT and treating depression, pain, anxiety, cancer and neurodegenerative
 PT diseases -
 PT
 XX Claim 6; Page 19-20; 75pp; English.
 PS
 XX The invention relates to a purified neutralised (Neu) polypeptide
 CC comprising at least one neutralised homology repeat (NHR) domain and a
 CC C3HC4 RING-zinc finger domain, and a polynucleotide encoding the
 CC polypeptide. The polynucleotide is useful for constructing a transformed
 CC host cell that expresses a Neu protein. Neu has neurogenic function and
 CC functions as a transcriptional regulator, as a calcium-signal transducer,
 CC in cell signalling and synaptogenesis, in memory learning,
 CC tumorigenesis, myogenesis and development of other organ systems and is
 CC related to repair and regeneration after injury to the central nervous
 CC system. Neu interacts with neurogenic genes, proteins implicated in
 CC nuclear transport and with Parkin-like proteins. Neu has the potential to
 CC interfere with inductive tissue interactions, cytokine signalling, RNA
 CC processing, early immediate responses, death of specific cell
 CC populations, nuclear hormone receptor signalling and axonal path-finding.
 CC The Neu family of proteins presents a set of diagnostic and therapeutic
 CC targets, especially for treating cancer. Manipulating Neu expression and
 CC function is useful in controlling a variety of diseases including
 CC depression, pain, anxiety, neurodegenerative diseases and cancer.
 CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
 CC invention.
 CC
 XX
 SQ Sequence 1725 BP; 268 A; 658 C; 520 G; 279 T; 0 other;

Query Match 24.4%; Score 409.2; DB 25; Length 1725;
 Best Local Similarity 56.6%; Pred. No. 1.4e-56;
 Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

Db 231 CCTCAGCCACAAAGGCTGTCAAGAGCAGAGCCAGCTTCTGCAACGCATCACTTCAGCAA 290
 Qy 219 GCGGCCATCTCGGCTGTACAGACAGTGCAGCTGCGCTGTGCGCCCTGCTG 278
 Db 291 CCGCCCGGCTCTCATCTACAGAGCAAGTCAAGCTCAAGATCAACAAAGAGTGTCTGCT 350
 Qy 279 GAGGAGGCGCTGTGCGCTTCCGCTTCAACCGCAGATCCGTGCTCAAGAGCCCAAGA 338
 Db 351 GAGCGAGGCTCTGCGGCTGAGCTTCAACAGCAAGACCCGTCCGATCACTGATCTC 410
 Qy 339 CATCCCAAGTACCGCTGCGCAGCTGTGTACCGCGCGGCTACTGAGCCAAAGCACT 398
 Db 411 GCTGCCAAGTACCGCTGCGCAGCTGTGTCCAGAGTGGCTTCTGAGCAAGGCGCT 470
 Qy 399 GCCCGAAGACTGTGCGCTGCGCAGACAGTGTGCTTCTGAGCCGACCGCACGCG 458
 Db 471 GCTGAGGAGTGTTCATACAGAGGCGCAATCATGCGATTCGTGGTGTGCAAGAGGCGCG 530
 Qy 459 GGTGTCTACAGCGGTGAACGAGCGGAGCGGCTCTTCACTGCGCGGTGCGCTGTGG 518
 Db 531 TGTCTTCAACGCAATCAACGACTCGCTGTATGTCTTTCAGCGGAGTCCGACGAGC 590
 Qy 519 CGGCCGCTCTGAGCGCTCATTTGATGTCTACGAGATCAACGACAGGTGCACTTGTGA 578
 Db 591 GCAACCGCTCTGAGCGCTGTGTGAGAGCTTACGCTCAACGCGGCGCTTCAGCTTGA 650
 Qy 579 GAGCGCTTGTGCTGTACAGCTGTACGCGCGCGGCTTCAACGAGCGCGCTTCAAGCGCTG 638
 Db 651 TAGCGAGCTGG-----TGTCCCGGACTGTGTGCGGCGCGCTTCAACGCGCTT 701
 Qy 639 CTTGCGCGCCAGACAGCAACGAGCGCAATTGTGACAAACAGAGCTGTGAAACAA 698
 Db 702 GCGGCGCGCTGTGCTGTGCGCGAGCGCAAGCGCGCTTCTGCGT----- 750
 Qy 699 GATGTGCGCAAGTGTGCGCACTGTGCGCTGTGCGCGCGCGCGCGCGCGCGCGCA 758
 Db 751 -----AGCTTATGTGCACTTCAAGTGTGCGCGCGCGCGCGCGCGCGCGCGCG 800
 Qy 759 GCGCGCGCGCGCGCGCATTTCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 818
 Db 801 GCT 860
 Qy 819 GCGCGTACTGTGAGCT 875
 Db 861 GCGCGAGTGTGAGCT 920
 Qy 876 GTCGCT 935
 Db 921 CTTGAG 980
 Qy 936 CTCGAGGCT 995
 Db 981 CACGAGCT 1040
 Qy 996 GCT 1055
 Db 1041 GCT 1100
 Qy 1056 CAAGAGGTGCT 1115
 Db 1101 GCGCGAGTGTGCT 1160
 Qy 1116 GCT 1175
 Db 1161 CCGCGTGTGCT 1220
 Qy 1176 GAGCGTGTCTGTGAGCATCAACGAGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCT 1235
 Db 1221 GAGGTGTACTGTGAGCAAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1280
 Qy 1236 GAGGCGCTGTGAGCGCTTCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1295

Db 1281 GCAGCGCTTTGATGCTCTTGGCCTGACGCGGACCATCACGAGATCCGATCCTCGG 1340
 QY 1296 CGGTACCTGAGTTCAGGCGCTTGCACGACATCCATCAGGGTCCCTCAG----- 1346
 Db 1341 CTCACATATCTGGCGGAGCGGGATATCCGATCATCCCTGCTCCCTGCTCCACGCC 1400
 QY 1347 -----CGGCTCCAGAGCATAGTATTCAGATATGACCTTCAGTTCACAGATC 1397
 Db 1401 AACCTGCGCCAGTGGCCCTGGGACGCGCTGTCTGACCCCTTCTCAGGACGATGAGCTC 1460
 QY 1398 CTCCTGGGATCTGATCATCTCTGTGACGCGCCCAAGTCTCCCGCTAGCCCCCGGT 1457
 Db 1461 TGGCCCTCTGGAGTCTCTGTGTGTGACAGCCCCCAATTCGCAGTAGCCTGCGCA 1520
 QY 1458 GTCCCGGCTTCTCCCAACCGGACCGGACGATCAAGAAATGGGAGTGCAGCGGTGTG 1517
 Db 1521 GTCCGAGTGAACCCAGGTTGGGCAATG-----GAGCATGATGTCACATTTG 1571
 QY 1518 CTTGATGGGAGTGAACAGGTCATCTACAGTGTGACACATGTGCTGTGCAAG 1577
 Db 1572 CATGAAACAGCGGTGGAACGCTCATCTACATGTGGCAATGTGCTGTGTAAGC 1631
 QY 1578 CTGCGGCTTGGGCTCAAGGACAGCGCGGCTGTGCTGCTCCATCTGCGGAGGCTCAT 1637
 Db 1632 CTGTGGCTTGGGCTCAAGAGGCTGTGACAGCTGTGCTGCTCCATCTGCGCGCCCAT 1691
 QY 1638 CAAGAGCTCATTAAGTCTACAG 1661
 Db 1692 CAAGAGATCATCAAGCTTACCG 1715

RESULT 9.

AAX77135
 ID AAX77135 standard; DNA; 2207 BP.

AC AAX77135;

DT 03-AUG-1999 (first entry)

DE Human nerve mutation factor protein encoding DNA.

KW Nerve mutation factor; chromosome 10; glioma; tumour suppressor;

KM brain tumour; astrocytoma; gene therapy; human; 88.

OS Homo sapiens.

PN MO925827-A1.

PD 27-MAY-1999.

PF 24-AUG-1998; 98MO-JP03737.

PR 14-NOV-1997; 97JP-0313211.

PA (SUME) SUMITOMO ELECTRIC IND CO.

PI Nakamura H, Nakata M, Saya H, Yoshida M;

DR WPI, 1999-347474/29.

DR P-PSDB; AAY21558.

PT Human gene on chromosome 10 homologous to Drosophila neuralized

PS Claim 9; Page 49-56; 78pp; Japanese.

CC The invention relates to a protein which is a nerve mutation factor and
 CC is the expression product of a gene located on chromosome 10. The gene
 CC is in a region frequently deleted in highly malignant gliomas. Sequences
 CC (AAX77135 and AAX77136) encoding human and mouse nerve mutation factors
 CC (AAY21558 and AAY21559) are provided. The protein is believed to have
 CC tumour suppressor activity. Polynucleotide sequences and antibodies to
 CC the protein are diagnostic reagents for highly malignant brain tumours

CC such as astrocytoma where chromosome 10 deletion commonly occurs. The
 CC gene may also be used for gene therapy of such tumours. The present
 CC sequence represents a DNA encoding the human nerve mutation factor.

SQ Sequence 2207 BP, 340 A; 872 C; 673 G; 322 T; 0 other;

Query Match 24.4%; Score 409.2; DB 20; Length 2207;
 Best Local Similarity 56.6%; Pred. No. 1.4e-56;
 Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

QY 99 GGTCTGGGAGAGGCGCGCTTTCACCGGAGCCAAAGCCAAAGATGCGGCTGA 158
 Db 581 GTCTCCAGACGCGCGCTGTCTTCAACCCGACACAAAGGCTCCCAATCTCATGGA 640
 QY 159 CGGCACTGCGCCCGGACACAGGCGCAACAGTTTCTGACATGGCGTCACTTTCACGA 218
 Db 641 CTTAGACCAAGAGCTGTCAAGAGGACGAGCCAGCTTTCAGAACGATCACTTCAAGAA 700
 QY 219 GCGGCCCATCCGACTGTACAGACAGTGCAGTGCCTGTGTGTGCGGCTGTGCTG 278
 Db 701 CGGCGCGGCTCATCTACAGCAAGTCAAGGCTGAAGATCACAAAGAGAGTGTGCTG 760
 QY 279 GAGCGGCGGCTGTGCGCTTTCACCGGAGACATCGGTGTCTATAGGCGCCAGGA 338
 Db 761 GAGCGGCGGCTGTGCGGCTTTCACCGAGAGACCGCTCGGATCACTGTGATC 820
 QY 339 CATCCCAAGTACGCTGCGCGGACCTGTGTACGCGGCGCGGCTACTGTGGGCAAGCACT 398
 Db 821 GCTGCCAAGTACGCTGTGCGCGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 880
 QY 399 GCGCGAGAACCTGGCGGTGCGGACACGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 458
 Db 881 GCCTAGAGAGTTTGCATATGAGAGGCAACATCTGATTTGTGTGTGTGTGTGTGTGTGT 940
 QY 459 CGT 518
 Db 941 TGTCTTTCACCGGATCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1000
 QY 519 CGGCGGCTGTGGGCGCTCATTTATGTCTACAGGACATCACAGACAGAGTGTGTGTGTGA 578
 Db 1001 CGACCGGCTGTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1060
 QY 579 GAGCGGCTTGT 638
 Db 1061 TAGCGAGCTGTG-----TGTCTCCGGAATGTGTGTGTGTGTGTGTGTGTGTGTGT 1111
 QY 639 CTTGCCGCGGACAGACAGACAGGCGGCACTTTCAGCAACAGAGTGTGTGTGTGTGA 698
 Db 1112 GCGGCGGCGGT 1160
 QY 699 GGT 758
 Db 1161 -----AGCTATGTGAGCTCAACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1210
 QY 759 CGCGCGGCGCGCGGCAATTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 818
 Db 1211 GCGCGCGGCGGT 1270
 QY 819 GCGGCTATGTGAGGCGGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 875
 Db 1271 GCGGAGCTGTGAGGAGGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1330
 QY 876 GTGCGGAGACCGCAAGT 935
 Db 1331 CTTGACGAGAGAGAGGT 1390
 QY 936 CTCGAGCGCGCGCTGTGCGGCGCGGAGAGGCTTCTGTGTGTGTGTGTGTGTGTGTGT 995
 Db 1391 CACCAAGCGCGCGCGGT 1450
 QY 996 GCGGCGGCGCGGCGGCGGT 1055
 Db 1451 CGGCGGCGCGGCGGCGGT 1510

QY 1056 CAACGAGCTGCCCCCGGACCCAGACGCGCTGCTGCACCGCAAGAGTACGTGGTGTGGC 1115
 DB 1511 GACCGACCTGCTTTCAGACCCCTGAGGCGCTGGTGAACCGAAGAAATTCGGGCGGTGTG 1570
 QY 1116 GCGGCGCGGCGCCGTCGAGACGGGGGAGACCGGCTCAGCTTCAAGCTTCGCGGCGGGGG 1175
 DB 1571 CCGGCTGCGCGGCGCCCTGACAGCGGCGACATCTGGGCGCTGGTGTCAACGCGGACGG 1630
 QY 1176 CGACGTGCTCTGGGACATCAACGGGCGGTCCGCGCGGCGCGCTGCTGTGCTGACACAC 1235
 DB 1631 CGAGCTGACCTCAGACCAATAGGCGCGCGCGCGCGCGCATGACGTGTGCTGAGAGCTTC 1690
 QY 1236 GCAGGCGCTCTGGGCGCTTCTTTCGCGCGGTGCGCGCGCGCTGCGGCGCGAGCTGCTCT 1295
 DB 1691 GCAGCGCTTGTGATGCTCTTTCGCGCGCTGACGCGGACATCAAGCAGATCCGATCTCGG 1750
 QY 1296 CGGTACCTGCAATGACCGCTGCGACCAACCATCTCAATGAGGTCCCTCAG----- 1346
 DB 1751 CTCCACTATCTGGCGCGGAGCGGGATATCCCGTCACTCCCTGCTCCCTGCTCCAGCGCC 1810
 QY 1347 -----CGGCTCCAGACGATGATGATTGATATGACCTTCACTGTCAACCAAGTTC 1397
 DB 1811 AACCTGCGCCAGTGGCTTGGGAGAGCGCGCTGTCTGACCCCTTGTCTCAGCAGTGAAGCTC 1870
 QY 1398 CTCTCGGCACTCTGAGTCACTCCGTGTGACGCGCGCGCGCGCTGAGCGCGCGCTG 1457
 DB 1871 TGGCCCTCTGGGTAGCTCTGTGTGTGAGACAGCGCCCAATTCGCGAGTGAAGCTGCCGA 1930
 QY 1458 GTCCCCCGCTTCTTCCCAACCGGAGCGCGGAGCATCAAGATGCGAGTGAACGCTGTG 1517
 DB 1931 GTCCGCAATGACCCAGGCTGTGGGCGAGTG-----GAGCGATGAGTGAACATTTG 1981
 QY 1518 CTTGATGCGGAGGTGACACAGGATCATCTACAGTGTGAGACATGTGCTGTGCGACAG 1577
 DB 1982 CTATATACACGCGGTGAGACAGGATCATCTACAGTGTGAGACATGTGCTGTGCGACG 2041
 QY 1578 CTGCGGCGCTGCGGCTCAAGCGACGAGCGCGCGCTGCTGCGCATCTGCGCGGCGCAT 1637
 DB 2042 CTGTGCGCTGCGGCTCAAGAGAGCTGTGACGCTGTGCGCATCTGCGCGGCGCGCAT 2101
 QY 1638 CAGGAGCTCATTAAGATCTACAG 1661
 DB 2102 CAGGAGCATCATCAAGACCTACCG 2125
 RESULT 10
 AAX77136
 ID AAX77136 standard; DNA; 1641 BP.
 AC AAX77136;
 XX
 DT 03-AUG-1999 (first entry)
 DE Mouse nerve mutation factor protein encoding DNA.
 XX
 KM Nerve mutation factor; chromosome 10; glioma; tumour suppressor;
 XX brain tumour; astrocytoma; gene therapy; mouse; ss.
 OS Mus sp.
 XX
 PN WO925827-A1.
 XX
 PD 27-MAY-1999.
 XX
 PF 24-AUG-1998; 98WO-JP03737.
 XX
 PR 14-NOV-1997; 97JP-0313211.
 XX
 PA (SUME) SUMITOMO ELECTRIC IND CO.
 XX
 FI Nakamura H, Nakata M, Saya H, Yoshida M;
 XX

DR WPI: 1999-347474/29.
 DR P-PSDB; AAY21559.
 PT Human gene on chromosome 10 homologous to Drosophila neuralized
 PT gene, useful in the diagnosis and gene therapy of brain tumors
 XX
 PS Claim 9; Page 61-66; 78pp; Japanese.
 CC The invention relates to a protein which is a nerve mutation factor and
 CC is the expression product of a gene located on chromosome 10. The gene
 CC is in a region frequently deleted in highly malignant gliomas. Sequences
 CC (AAX77135 and AAY21559) encoding human and mouse nerve mutation factors
 CC (AAY21558 and AAY21559) are provided. The protein is believed to have
 CC tumour suppressor activity. Polynucleotide sequences and antibodies to
 CC the protein are diagnostic reagents for highly malignant brain tumors
 CC such as astrocytoma where chromosome 10 deletion commonly occurs. The
 CC gene may also be used for gene therapy of such tumors. The present
 CC sequence represents a DNA encoding the mouse nerve mutation factor.
 XX
 SQ Sequence 1641 BP; 260 A; 594 C; 498 G; 288 T; 1 other;
 Query Match 23.4%; Score 392.2; DB 20; Length 1641;
 Best Local Similarity 55.9%; Pred. No. 76-54;
 Matches 885; Conservative 0; Mismatches 639; Indels 60; Gaps 5;
 QY 99 GATCTGTGGCGAGGCGCGCGCTTTCACGCGCGCAGCGCAAGAAAGTGGCGCTGGA 158
 DB 87 GCTCCCGGCGACGCGCGCTGCTTTCACCGCGCGCATTAAGGCTCCCGAGATCTCATGGA 146
 QY 159 CGGCACTCGGCG 218
 DB 147 CTTAGCGACAGAGCG 206
 QY 219 GCGGCGCGATCCGCGCTGACAGAGAGTGGCGCTGCGCGCGCGCGCGCGCGCGCGCG 278
 DB 207 CCGCGCGGCTGCTATCAAGAGAGAGTGAAGTCAAGTCAAGAGAGAGTGTGCTG 266
 QY 279 GAGCGGCGCGCTGCGCTTTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 338
 DB 267 GAGCGGCGCGCTGCG 326
 QY 339 CATCCCGAAGTACCG 398
 DB 327 GCTGCGCGAGTGGCG 386
 QY 399 GCCGAGAACTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 458
 DB 387 GCTGAGGAGTGTGCG 446
 QY 459 CGTGTCTACAGCGTGAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 518
 DB 447 GGTCTTCAACGAGTCAATGAGTCAAGCTGCTATGCTTTCTTCAAGTGGGATCCGACGGT 506
 QY 519 CGGCGCGGCTGCG 578
 DB 507 GACCGCGCTGCG 566
 QY 579 GAGCGCTTTCGCTGACAGCGTGAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 638
 DB 567 CAGCGAGGTGG-----TGTGCGCGGACGTGCTGCGCGCGCGCGCGCGCGCGCGCT 617
 QY 639 CTTGCG 698
 DB 618 GCGGCG 656
 QY 699 GGTGTGCGCAAGCTGCG 758
 DB 657 CCTGTGCGTGAAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 716
 QY 759 CG 818
 DB 717 GCTGCG 776

QY 819 GAGCGTACTGAGAGCGGACCTGCGCTTCCACGCAACCGCGGCGCG--ACGTGAGCCT 875
 DB 777 GGGCGAGCTGACGCGGACCTGCGCTTCCACGCGGCTTCCGCGCGCGCGACGCTCGGCAT 836
 QY 876 GTGGCGGACCGGCAAGAGTGGCTGCGGACCGCGCGCGCGCGCGCGCGCGCTT 935
 DB 837 CTTGAGACGACGAGCGGTGGCGCGCTGAGACCGGCGCGACGCGCGCTCGCTT 896
 QY 936 CTCGAGCGCGCGCGCTGCGCGCGCGCGCGAGCGCTTCTGAGAGTGGCGCGCGCGCT 995
 DB 897 CACGAGCGCGCGCTGAGCGCTGCGCGCGAGACCATCTTTCATCAAGGTACAGCGCTCGCGG 956
 QY 996 GCGCGCGCGCGCGCGCTGCGCTTCCGCTTCCGCTGCGAGCGCGCGCGCTGCTTACGCGC 1055
 DB 957 GCGCGCGAGCGCGCGCTGCTTCCGCGGTCCACGACGTGAGACCTGAGCGCGCGCGCC 1016
 QY 1056 CAACGAGCTGCGCGCGCGCGCGCGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCG 1115
 DB 1017 CCGCGAGCTGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1076
 QY 1116 GCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCGCTGCGCGCGCG 1175
 DB 1077 TCCGCTGCGCGCGCGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1136
 QY 1176 CGAGCTGCTCTGCGCGCATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGCGCGCG 1235
 DB 1137 AGAGCTGACCTGAGTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1196
 QY 1236 GCGAGCGCTGCGCGCGCTTCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1295
 DB 1197 GCGCGCGCTGCGCGCGCTTCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1256
 QY 1296 CGGTAACCTGCGAGTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1348
 DB 1257 CTCGACCATATGATGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1316
 QY 1349 -----GCTCCAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1397
 DB 1317 AACCTGACCGAGTCCG 1376
 QY 1398 CTCCTGCGCATGAGTATCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1457
 DB 1377 TGGCGCGCTGAGTGGCTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1436
 QY 1458 GTCCCGCGGTCTTCCCG 1517
 DB 1437 GCCACCGGTACCCAGGTCTGGCGCG-----TGAGTATGATGATGATGATGATGAT 1487
 QY 1518 CTTGATGCGGAGGTGACACGCGTATCTACGCTGAGACACATGCTGCTGCTGCTGCTG 1577
 DB 1488 CTATGAAACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1547
 QY 1578 CTGCGCGCTGCGCGCTGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1637
 DB 1548 CTGCGCGCTGCGCGCTGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1607
 QY 1638 CAAGGACGTATTAAGTCTACAG 1661
 DB 1608 CAAGGACATATCAAGACTTACG 1631

RESULT 11
 ABX75852
 ID ABX75852 standard; cDNA; 1698 BP.

XX ABX75852;

DT 31-MAR-2003 (first entry)

XX Mouse Neu1 cDNA #2.

KM Human; mouse; rat; gene; 86; neutralised polypeptide; Neu; NHR; cancer;
 neutralised homology repeat; C3HC4 RING-zinc finger domain; myogenesis;

KM neurogenic function, transcriptional regulator, cell signalling, pain;
 KM calcium-signal transducer; synaptogenesis; memory learning; anxiety;
 KM tumorigenesis; organ development; central nervous system; depression;
 KM neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;
 KM inductive tissue interaction; cytokine signalling; axonal path-finding;
 KM neurodegenerative diseases; cytostatic; antidepressant; analgesic;
 KM tranquiliser; neuroprotective; Neu1; Neu2; Neu3; Neu4.
 XX
 XX Mue sp.
 OS
 PN US2002132293-A1.
 XX
 XX 19-SEP-2002.
 XX
 XX 14-MAR-2001; 2001US-0808387.
 XX
 XX 14-MAR-2001; 2001US-0808387.
 XX
 XX (PALM/) PALM K.
 PA (TIMM/) TIMMUSK T.
 PI Palm K, Timmusk T,
 XX
 XX DR WPI, 2003-174081/17.
 DR P-PSDB, ABUS6245.
 XX
 XX Novel neutralized polypeptide, Neu useful for controlling cell
 PT proliferation and calcium signaling induced transcriptional processes
 PT and treating depression, pain, anxiety, cancer and neurodegenerative
 PT diseases -
 XX
 XX PS Claim 6, Page 28-29; 75pp; English.
 CC The invention relates to a purified neutralised (Neu) polypeptide
 CC comprising at least one neutralised homology repeat (NHR) domain and a
 CC C3HC4 RING-zinc finger domain, and a polynucleotide encoding the
 CC polypeptide. The polynucleotide is useful for constructing a transformed
 CC host cell that expresses a Neu protein. Neu has neurogenic function and
 CC functions as a transcriptional regulator, as a calcium-signal transducer,
 CC in cell signalling and synaptogenesis, in memory learning,
 CC tumorigenesis, myogenesis and development of other organ systems and is
 CC related to repair and regeneration after injury to the central nervous
 CC system. Neu interacts with neurogenic genes, proteins implicated in
 CC nuclear transport and with Parkin-like proteins. Neu has the potential to
 CC interfere with inductive tissue interactions, cytokine signalling, RNA
 CC processing, early immediate responses, death of specific cell
 CC populations, nuclear hormone receptor signalling and axonal path-finding.
 CC The Neu family of proteins presents a set of diagnostic and therapeutic
 CC targets, especially for treating cancer. Manipulating Neu expression and
 CC function is useful in controlling a variety of diseases including
 CC depression, pain, anxiety, neurodegenerative diseases and cancer.
 CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
 CC invention.
 CC
 XX
 XX Sequence 1698 BP; 277 A; 608 C; 514 G; 299 T; 0 other;
 SQ
 Query Match 23.4%; Score 391.6; DB 25; Length 1698;
 Best Local Similarity 55.9%; Pred. No. 8.7e-54;
 Matches 885; Conservative 0; Mismatches 639; Indels 60; Gaps 5;

Db 300 GAGCGGGGCGCTGCGCACTTGGCTTCAACGACGAGAACCCCTTCCCGATCAACCCCACTC 359

Qy 339 CATCCCCAAGTACGCGCTGCGCGGACCTTGGTCAACGCGGCGGGGCTACTGGGCGCAAGGCACT 398

Db 360 GCTGCCAAGTACGCGCTGCGCTGACCTGGTGTCTGAGAGTGGCTTCTGGGCGCAAGACTT 419

Qy 399 GCCCGAAGAACTTGGCGCTGCGGACACAGGTGCTGGCTTACTGGGCGGACCGCACGCGCG 458

Db 420 GCGTAGAGAGTTTGGCAACGAGGGCAACATCATTTGCTTCTGGTGTGAACAAGAGGCGG 479

Qy 459 CGTGTCTTACAGCGTGAAGACAGCGGCGAGCGGGGCTTCACTGCGGGGCGGTGGG 518

Db 480 GGTCTTCTTACCGGAATCAATGAAGTCAAGCTGTATGCTTTCTTCAAGTGGGGTCCGACGGT 539

Qy 519 CGGCGCGGCTTGGGCGCTCATTTGATGTCTTACGCGCATCAACGACGAGGTGCACTTCTTGA 578

Db 540 GGACCCGCGCTCTGGGCGCTGTGTGAAGTCAACGCGCTCAAGCGGGGTGTCAAGCTGTAGA 599

Qy 579 GAGCGCGCTTCCGTGAACGCTGACGCGCGCGCGGCTCAGCGACGCGCGCTTACGCGCTG 638

Db 600 CAGCGAGCTGG-----TGCTGCGCGACATGCTCGGAGCGCGCTCTTCAACCGCGCT 650

Qy 639 CCGCGCGCGCCAGACGCGCAACGAGCGCGCGCACTTCCAGCAACAGAGCTGAGAACAAACA 698

Db 651 GCGGCGCGCGCTGCGCTGCGGTGCGAGGCG-----GATGAGGCGCG 689

Qy 699 GGTGTGTGCGCAAGCTTGGGCAACTTGGCGCTGTGGGCGCGCGCGCGCGCGCGCAACCGCAACG 758

Db 690 CCGTGTGCGTGAAGCTGTGTGACCTCAACGTGCGGGAGCGAGCGAGCGAGACGAGCGCGCAC 749

Qy 759 GCGCGCGGCGCGCGCATTTCCGTGTGGGAGCGCGCGGTGAAGCGCGCGCGCGCGCGCTGTGCTC 818

Db 750 GCCTTCCGAGCTCCCGATCCCGACAGACTCCGCTCAATTCAGACAGAGCGCGCGCTGGC 809

Qy 819 GCGGCTACTGAGAGCGCGACCTTGGCTTCCAGCGCAACGCGGGCGCG---AGTGAAGCT 875

Db 810 GCGCGAGCTCGACGCGCGACCTTGGCTTCCAGCGCTTGGCGCGCGCGCGCGCAAGTCCGCAAT 869

Qy 876 GTGCGCGCGACCGCAAGTGTGCGCTGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 935

Db 870 CCTGGAACGAGACGAGTGTGGCGCGCTTGGAGACACGGGCGCGAGACGAGCGCGCTGCTT 929

Qy 936 CTCCGAGCGCGCGCTGCGCGCGCGCGCGAGCGCTTCTGTGTGAAGTGGGCGCTTCCGAGGCT 995

Db 930 CACCAAGCGCGCGCTGTGAGCGGTGGCGAGCAATCTTCAATGAAGTCAACGCGCGTCCGGCGG 989

Qy 996 GCGCGCGCGCGCGCGCTTGGCATCAAGTGTGCGACCGCGCGCGTGTCAAGGCG 1055

Db 990 GGGGAGACGCGCGCGCTGTCTTCCGAGGTCAACAGTGTGACCTTGGACGCGTGGCGCGC 1044

Qy 1056 CAACGAGCTGCGCGCGCGCAACCGAGCGCGCTGCTCGACCGCAAGAGTACCTGGGTGGGCG 1111

Db 1050 GCGGACCTGCGCTTCAAGCTCCGAGGCGCTTGGTGAACCGAAGAACTTCTGGCGGGTGTG 1100

Qy 1116 GCGCGCGCGGCGCGCTGCGCGAGCGCGCGAGCGCGCTCAAGCTTCAACGCTGCGCGCGCGG 1177

Db 1110 TCGCGTGGCGCGGCGCTGTGCAACAGGGGCGCAATCTCGGGCGTGTGTGTCAACGCGGACGG 1166

Qy 1176 CGACGTGCTCTTGGCATCAACGGCGGTCCGCGCGCGCGCTGTGTGTGTGACACAC 1233

Db 1170 AGAGCTGCAACTGAGTCAACAAAGCGCGCGCGCGCGCGCAAGCTGTGTGTGATGCTTC 1222

Qy 1236 GACGCGCGCTTGGGCGCTTGTGCGCGTGTGGCGCGCGCGCGTGTGGGCGCACTGTGTGCTC 1299

Db 1230 GGAAGCGCGCTTGTGAAGCTTCAAGCTGTGAAGCGCAATCAACGAGTCCGATCTTCGG 1288

Qy 1296 CGGTACCTTGAAGTCAAGCGCTTGCACACGACTCCATCAAGGATCC----- 1344

Db 1290 CTCACACATCAAGACTGAACGGGGGTGGCGCAATCTCCCTGTGTACCTGTTCACCTCC 1344

Qy 1343 -----TACGCGGCTCCCAAGACGATGTGATTGACTTCAAGTGTCAACGAGTC 1399

Db	1350	AACTCAACCCAGTGCCTCGGGGACACCGCTTGTCGACCCCTTGTCAGACCTGCGGGTTC	1409
Qy	1398	CTCCTCGGCACTGTAGTCATCTCCCTGTGTGACGAGCCGCCAGCTCCCGCTGAGCCCGCGGT	1457
Db	1410	TGGGCCCCCTAGTGTGACTCTGTGTGAGGAGACAGCCCCCAATTCACCTGTGAGCCCTGCCGA	1469
Qy	1458	GTCCCCCGTGTCTTCCCAACCGGAGCCGGCAGGCGATCAAGATGAGCGAGTGCACGGTGTG	1517
Db	1470	GTCAACCGGTGACCCCAAGTCTGGGCCAG-----TGAAGTATGAAATGCACCATTTG	1520
Qy	1518	CTTGGAATGCGGAGGTGACACGGTCACTCAAGTGTGTGACACATGTGTGCTGTGTGCCACAG	1577
Db	1521	CTATGAACACCGAGTGTGATACAGTCACTTACAGTGTGGCCACATGTGCTGTGTCTACTC	1580
Qy	1578	CTGCGGCGCTGTGGGCTCAAGCGACAGGCCCCGGGCGTGTGCCCATCTGTGCCGCGGCCAT	1637
Db	1581	CTGTGGCGCTGTGGCGCTCAAGAAAGCGCCCTGTGCACGCGCTGTGCCCCCATCTGCGGCCAT	1640
Qy	1638	CAAGGAGCTCATTAAGATCTTACAG	1661
Db	1641	CAGGACATCATCAAGACTTACCG	1664

XX	ABX75851	standard; cDNA; 1749 BP.
XX	ABX75851,	
XX	31-MAR-2003	(first entry)
XX	Mouse Neu1 cDNA #1.	
XX	Human; mouse; rat; gene; 89; neutralised polypeptide; Neu; NHR; cancer;	
XX	neutralised homology repeat; C3HC4 RING-zinc finger domain; myogenesis;	
XX	neurogenic function; transcriptional regulator; cell signalling; pain;	
XX	calcium-signal transducer; synaptogenesis; memory learning; anxiety;	
XX	tumorigenesis; organ development; central nervous system; depression;	
XX	neurogenic gene; nuclear transport; Parkin-like protein; RNA processing;	
XX	inductive tissue interaction; cytokine signalling; axonal path-finding;	
XX	neurodegenerative disease; cytoskeletal; antidepressant; analgesic;	
XX	transferrin; neuroprotective; Neu1; Neu2; Neu3; Neu4.	
XX	Mus sp.	
XX	US2002132293-A1.	
XX	19-SEP-2002.	
XX	14-MAR-2001; 2001US-0808387.	
XX	14-MAR-2001; 2001US-0808387.	
XX	(PALM/) PALM K.	
XX	(TIMM/) TIMMUSK T.	
XX	Palm K, Timmusk T;	
XX	WPI; 2003-174081/17.	
XX	P-PSDB; AB056244.	
XX	Novel neutralized polypeptide, Neu useful for controlling cell	
XX	proliferation and calcium signaling induced transcriptional processes	
XX	and treating depression, pain, anxiety, cancer and neurodegenerative	
XX	diseases -	
XX	Claim 6; Page 25-26; 75pp; English.	
XX	The invention relates to a purified neutralised (Neu) polypeptide	
XX	comprising at least one neutralised homology repeat (NHR) domain and a	
XX	C3HC4 RING-zinc finger domain, and a polynucleotide encoding the	
XX	polypeptide. The polynucleotide is useful for constructing a transformed	
XX	host cell that expresses a Neu protein. Neu has neurogenic function and	

CC functions as a transcriptional regulator, as a calcium-signal transducer,
 CC in cell signalling and synaptogenesis, in memory learning,
 CC tumorigenesis, myogenesis and development of other organ systems and is
 CC related to repair and regeneration after injury to the central nervous
 CC system. Neu interacts with neurogenic genes, proteins implicated in
 CC nuclear transport and with Parkin-like proteins. Neu has the potential to
 CC interfere with inductive tissue interactions, cytokine signalling, RNA
 CC processing, early immediate responses, death of specific cell
 CC populations, nuclear hormone receptor signalling and axonal path-finding.
 CC The Neu family of proteins presents a set of diagnostic and therapeutic
 CC targets, especially for treating cancer. Manipulating Neu expression and
 CC function is useful in controlling a variety of diseases including
 CC depression, pain, anxiety, neurodegenerative diseases and cancer.
 CC Sequences ABX75848-ABX75864 represent Neu polynucleotides of the
 CC invention.

SQ Sequence 1749 BP; 285 A; 630 C; 526 G; 308 T; 0 other;

Query Match 23.4%; Score 391.6; DB 25; Length 1749;

Best Local Similarity 55.9%; Pred. No. 8.6e-54;

Matches 885; Conservative 0; Mismatches 639; Indels 60; Gaps 5;

QY 99 GGTCTGGAGGAGGCGCGCTTCCAGCGCAAGGCCAAGCAAGAACTGCGGCTGGA 158
 DB 171 GCTCCGGGCAACCGCGCTCTTCCACCCCACTAAGGGCTCCAGATCTCTAATGA 230
 QY 159 CGGCACTCGCGCGCGGCAACAGGCGCAACAGCTTCAATAGGCGTCAAGTCAAGCA 218
 DB 231 CTTAAGCACAGAGCGCTCAAGAGGCAAGGCGCAAGCTTCAATAGCACTTCAAGTAA 290
 QY 219 GCGGCCATCCGCGCTGACAGAGAGTGGCGCTGCGCGCTGAGCGCGCTGAGCTG 278
 DB 291 CGCGCGGCTCATCTACAGAGCAAGTCAAGCTGAAGATCAACAAAGAAATGCGCTG 350
 QY 279 GAGCGCGCGCTGCGCTTCCAGCGCAAGATCGCTCATGAGCGCCAGGA 338
 DB 351 GAGCGGGGCGCTGCGCTTCCAGCGCAAGATCGCTTCCCGCATCCACCGGCTC 410
 QY 339 CATCCCAAGTAAAGCGCGCGGCAAGCTGACAGCGCGCGGCGCTAAGGCGCAAGGACT 398
 DB 411 GTCGCCAAGTAAAGCGCGCGCTGACAGCGCGCGGCGCTTCCGCGCAAGGACT 470
 QY 399 GCCCGAAGACTGAGCGCTGCGGCAACAGTGTGCGCTTCAAGCGCGCAAGCGCG 458
 DB 471 GCTGAGAGATTTGGCCAAAGAGGCAATATGCTTCTGAGTGAACAAGAGCGCG 530
 QY 459 CGTGTCTACAGCGTGAACAGCGCGAGCGGCTCTTCACTGCGCGCTGCGCTGAG 518
 DB 531 CGTCTCTACCGGATCAATGATGATGCTGCTATGCTTTCTTCAATGAGGCTCCGAG 590
 QY 519 CGCGCGCGCTGCGCGCTGCAATGATGATGCTGCAAGGATGAGGCTTCTGGA 578
 DB 591 GAGACCGCTGCGCGCTGAGTGTACGCGCTTCAACGCGGAGTGTCAAGCTGCTAGA 650
 QY 579 GAGCGCGCTTCTGCTGACAGCTGACGCGCGCGCGCTGAGCGCGCGCTTCAAGCGCTG 638
 DB 651 CAGCGAGCTGG-----TGTGCGCGGAGCTGCGCGCGCGCTTCTTCAACGCGGCT 701
 QY 639 CTGCGCGCGCAAGCAAGCAAGCGCGCAACTTTCAGCAACAAGAGCTCAAGAAACA 698
 DB 702 GCGGCGCGCGCTGCGCTGAGTGTGAGGCG-----GATGAGGCGCG 740
 QY 699 GGTGTGCGCAAGGTGAGGCGCAAGTGTGAGGCGCGCGCGCGCGCGCGCGCGCAAG 758
 DB 741 CTTGTGAGTGAAGCTTGTGCACTTCAAGTGTGCGGAGCGCAAGCGCGCGCAAGC 800
 QY 759 CGCGCGCGCGCGCAATTCGTGCGGAGCGCGCGCGCGCGCGCGCGCGCTGCTGCGC 818
 DB 801 GCGTGGCGCGCGCGCGCAATTCGTGCGGAGCTGCTCAATTCAGAGCAAGCGCGCGCTG 860
 QY 819 GCGGCTACTGAGGCGCGCACTGCGCTTCAAGCAAGCGCGGCGCG---ACGTGAGCTT 875
 DB 861 GCGGCACTGAGCGCGCACTGCGCTTCAAGCGCGCTTCTGCGCGCGCGCGCGCACTGCGCAT 920

QY 876 GTTGGCGCAACCGCAAGTGTGCTGTGCAACCGCGCGCGCAAGCGCGCGCGCACTGCTT 935
 DB 921 CTTGAGCAAGAGAGAGAGTGGCGCGCTGTGAGAGCAAGCGCGCGCAAGCGCGCGCTT 980
 QY 936 CTCCAGAGCGCGCTGTGCGCGCGCGCGAGAGCTTCTTGTGAGTGTGAGCGCGCT 995
 DB 981 CACAGAGCGCGCTGTGAGAGTGTGCGAGAGCAATCTTCAATCAAGTGTGACCGCGCGCG 1040
 QY 996 GCGCGCGCGCGCGCGCGCTTGTGAGTGTGCAAGTGTGAGCGCGCGCGCGCTACGCGC 1055
 DB 1041 GAGGAGAGAGAGAGAGAGTGTGCTTGTGAGGTGACAGAGTGTGACCTGTGAGCGCGC 1100
 QY 1056 CAAGAGTGTGCGCGCGCAACCAAGAGCGCTGTGCAACCGCAAGAGTATCTGTGTGTGCG 1115
 DB 1101 CGGAGACTGCGCTTCAAGCGCGCGCGCGCTGTGTGAGCGCAAGAGTGTGAGCGCGCTG 1160
 QY 1116 GCGCGCGCGCGCGCGCGCGCGCGAGCGCGCGCGCGCTCAAGCTTCAAGCTGCGCGCG 1175
 DB 1161 TCGCGTGTGCGCGCGCTGTGCAAGCGCGCGCAATCTGTGCGCTGTGTGTCAAGCGCGAG 1220
 QY 1176 CGAGTGTCTGTGAGCATCAAGCGCGCGCGCGCGCGCGCGCTGTGTGTGTGCAACAGC 1235
 DB 1221 AAGAGTGTCACTGAGTCAACAGCGCGCGCGCGCGCGCGCATGTGTGTGTGAGTGTCTC 1280
 QY 1236 GCAAGCGCTGTGCGCGCTTCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 1295
 DB 1281 GCAAGCGCGCTGTGAGTGTCTTCAAGCGCGCGCGCGCGCGCGCGCATGTGTGTGTGAG 1340
 QY 1296 CGGTACCTGTGAGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1342
 DB 1341 CTTCACATCAATGATGAGAAAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1400
 QY 1343 -----TCAAGCG 1397
 DB 1401 AACCTGACCAAGTGTGCG 1460
 QY 1398 CTCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1457
 DB 1461 TGGAGCG 1520
 QY 1458 GTTCCCGCGTGTGTCTCCCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1517
 DB 1521 GTCAACCGTGTGAGCG 1571
 QY 1518 CTTCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1577
 DB 1572 CTATGAAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1631
 QY 1578 CTGCGCGCTGTGCGCGCTTCAAGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1637
 DB 1632 CTGTGTGCTGTGCGCTTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1691
 QY 1638 CAAGAGCTGATTAAGATCTACAG 1661
 DB 1692 CAAGAGCATATCAAGAGCTACCG 1715

RESULT 13

ABX75855 standard; cDNA, 1725 BP.

ABX75855;

31-MAR-2003 (first entry)

Rat Neu1 cDNA #1.

Human; mouse; rat; gene; 88; neurallised polypeptide; Neu; NHR; cancer;
 neuronalised homology repeat; C3H4 RING-zinc finger domain; myogenesis;
 neurogenic function; transcriptional regulator; cell signalling; pain;
 calcium-signal transducer; synaptogenesis; memory learning; anxiety;
 tumorigenesis; organ development; central nervous system; depression;

Db 1395 CACTCCAACTGCGCCAGTGCCTTGGGAGACCGCCTCTCTGACCCCTGCTCAGACATG 1454
 Qy 1392 CCACTCTCTCTGCGCATCTGATGATCCCTGATGACGCGCCCAAGCTCCCGCTGAGCC 1451
 Db 1455 CCGCTTGGGCGCCCTGGGTGGCTCTGTGGCGGAAAGCCCACTCAGCTGTGAGCCT 1514
 Qy 1452 CCGGTGTGTCCTGCTGTCTCCCAACCGGACCGGCGGATCAAGATGAGCGAGAC 1511
 Db 1515 GCCCGAGTACAGTACCCCGGCTG-----GCCAGTGAAGCAATGAATGAC 1565
 Qy 1512 GGTGTCTTCTGATGGGAGTGAACAGCTCATCTACGCTGTGACACATGTGCTGTG 1571
 Db 1566 CATTGCTATGAACACGCAAGTGAATCAGTCACTACGCTGTGGCACAATGTGCTGTG 1625
 Qy 1572 CCACTGCTGGGCTGCGGCTTCAAGGACAGCGCCGCGCTGCTGCTGCTGCGGCG 1631
 Db 1626 CTACTCTGTGGCTGCGCTCAAGAAAGCCCTGACGCTGTGCTGCTGCTGCTG 1685
 Qy 1632 GCCCATCAAGACGTCTTAAGATCTACAG 1661
 Db 1686 CCCCATCAAGACATCTCAAGACTACCG 1715
 RESULT 14
 ABO43990
 ID ABO43990 standard; DNA; 1016 BP.
 AC ABO43990;
 XX
 DT 12-JUL-2002 (first entry)
 XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30581.
 XX
 KW Human; cytosine methylation; 5'-CpG-3', uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-BP10074.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 PA (EPIC-) EPIDENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 DR WPI; 2002-371829/40.
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA.
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridization to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridized to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (1) for diagnosis and/or prognosis of side effects of

CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNPs); and (11) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABO43910-ABO454121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX

SO Sequence 1016 BP; 123 A; 154 C; 337 G; 399 T; 3 other;

Query Match 22.7%; Score 380.6; DB 24; Length 1016;

Best Local Similarity 72.6%; Pred. No. 56-52;

Matches 530; Conservative 0; Mismatches 197; Indels 3; Gaps 3;

Qy 571 CTTCTGGAAGAGCGCTTGTGCTGACAGCTGAGCGCCGCGGCTGACGAGCGCGCTTC 630
 Db 98 CGATAGAGAGCGTTTCGTATACG-TGACGTTGCGCGTTTATGTTAGTTTCGTTT 156
 Qy 631 AGCGCTGCTGCTCGCCGACAGCAGACGCGCGCCACTTGACAAACAGAGCTCGAG 690
 Db 157 AGCGTTGTTTGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 216
 Qy 691 AACACACAGTGTGCGCAAGCTGCGCACTGCGCTGCGCGCGCGCGCGCGCGCGCG 750
 Db 217 AATTAATTAGTGTGTGTATGTTAGTTGTTATTCG-GTTGCGCTGCTTTCGTTTATG 275
 Qy 751 CCAAGCGAGCG 810
 Db 276 TTATGACAGCTGCGCGCTGCTGTTATTTGTTGTTGTTGTTGTTGTTGTTGTTG 335
 Qy 811 TCGTCGCGCGCTACTGAGAGCGCAGCTGCTGCTTCAAGCAACAGCGCGCGCGCGCG 870
 Db 336 TCGTGTGCGCGCTATGAGAGTCAATTTGCGTTTATGATGATGAGCGCGTTGCGAG 395
 Qy 871 AGCTGTGCGCGCGCAAGTGTGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 930
 Db 396 AATTGTGCGTCAATGTAAGTGTGCTGATGATGATGATGATGATGATGATGATGAT 455
 Qy 931 GTCCTTCTGAG-CGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 989
 Db 456 GTTTTTCGAGATTTTCGTGCGTACGAGATTTTTCGTGAGAGTGTGCTGCTG 515
 Qy 990 GGGGCTGCGCGCGCGCGCGCGCGCTTCCGATCAGCTGTGCGAACCAGCGCTGCT 1049
 Db 516 GGGGTTGGCGCGCTTGGCGCGCTTGGCTTTCGATTAACGTCGTGCGATTCGCGCG 575
 Qy 1050 ACGGCGCAAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1109
 Db 576 ACGGTTTAAAGAGTGTGCTGATTTAAGCGCGTGTGCTGATCTTAAAGAGTGTG 635
 Qy 1110 GGTGGCG 1169
 Db 636 GGTGGCG 695
 Qy 1170 CCGCGCGCGAGTGTCTCTGCGCATCAACGCGCGCTCGCGCGCGCGCTGTGCGTCA 1229
 Db 696 CCGCGCGCGAGTGTGTTTGGGATTAACGCGCGCTTGGCGCGCTTGTGTGCGTCA 755
 Qy 1230 CACCAAGCAGCGCTCTGCGCTTCTTCCGCGTGCAGCGCGCGCGCGCGCGCGCG 1289
 Db 756 TATTAGTATAGCGCTTTTGAGGTTTTCGTCGCGCGCGCGCGCGCGCGCGCGCG 815
 Qy 1290 TCTCTCTCGAT 1299
 Db 816 TTTTTCGAT 825
 RESULT 15
 ABO43991/C
 ID ABO43991 standard; DNA; 1016 BP.
 XX

AC AB043991;
 XX 12-JUL-2002 (first entry)
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30582.
 XX
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KM drug; side effect; cancer; central nervous system; cardiovascular;
 KM gastrointestinal; respiratory system; single nucleotide polymorphism;
 XX SNP; cell differentiation; de.
 XX
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-EP10074.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 PA (EPIC-) EPIDEMIOLOGICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX
 DR WPI; 2002-371829/40.
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 CC
 XX
 SQ Sequence 1016 BP; 399 A; 337 C; 154 G; 123 T; 3 other;

Query Match 22.7%; Score 380.6; DB 24; Length 1016;

Best Local Similarity 72.6%; Pred. No. 5e-52; Indels 3; Gaps 3;

Matches 530; Conservative 0; Mismatches 197;

QY 571 CTTCTGAGAGCGCTTCGCTGACGCTGACGCGCGCGCTTCAGCAGGCGCGCTTC 630
 DB 919 CGTATAGAGCGCTTCGCTGATAGG-TGACGCTTCGCGCTTCAGTTCGCTTC 861
 QY 631 AGCGCTGCTGCG 690
 DB 860 AGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801
 QY 691 AACACACAGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 750
 DB 800 AATTAATTAGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 742

QY 751 CAGCGCGAGCG 810
 DB 741 TTATCGAGCTGCG 682
 QY 811 TCGTCG 870
 DB 681 TCGTCG 622
 QY 871 AGCGCTGCG 930
 DB 621 AATTGTCG 562
 QY 931 GTCCTTCG 989
 DB 561 GTTTTTCGAGATTTTCGTTGCGGTGACGAGATTTTTCGTTGAGAGGTTCG 502
 QY 990 GGGCTGCG 1049
 DB 501 GGGCTGCG 442
 QY 1050 ACGGCG 1109
 DB 441 ACGGTTTAAAGAGTTGTTGCGATTTAGACGCGTTGTTGATCGTAAAGAGT 382
 QY 1110 GGTGCG 1169
 DB 381 GGTGCG 322
 QY 1170 CGGCG 1229
 DB 321 CGGCG 262
 QY 1230 CACGAGCG 1289
 DB 261 TATTACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 202
 QY 1280 TCTCTCTCGGT 1299
 DB 201 TTTTTCGCGT 192

Search completed: January 20, 2004, 15:21:47
 Job time : 468 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 20, 2004, 17:03:09 ; Search time 571 Seconds
(without alignments)
10339.717 Million cell updates/sec

Title: US-09-808-387-21

Percent score: 1675
Sequence: 1 atggcgacacacgctgcacgc.....ctacagcgcatagctcagcc 1675

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1675	100.0	1675	10	US-09-808-387-21
2	1228	73.3	1641	10	US-09-808-387-27
3	1103.6	65.9	1129	10	US-09-808-387-23
4	577.2	34.5	955	10	US-09-808-387-25
5	409.2	24.4	1674	10	US-09-808-387-1
6	409.2	24.4	1725	10	US-09-808-387-1
7	391.6	23.4	1698	10	US-09-808-387-9
8	391.6	23.4	1749	10	US-09-808-387-7
9	380.8	22.7	1725	10	US-09-808-387-15
10	264	15.8	1161	10	US-09-808-387-5
11	186.6	11.1	694	13	US-10-029-386-22715
12	182.4	10.9	876	10	US-09-808-387-13
13	177	10.6	888	10	US-09-808-387-19
14	176.8	10.6	1035	10	US-09-808-387-17
15	176.4	10.5	1035	10	US-09-808-387-11

16	166.4	9.9	789	10	US-09-808-387-29	Sequence 29, Appl
17	135.8	8.1	765	10	US-09-808-387-31	Sequence 31, Appl
18	127.8	7.6	693	13	US-10-027-633-152858	Sequence 152858
19	127.8	7.6	693	14	US-10-027-633-152858	Sequence 152858
20	119	7.1	536	13	US-10-029-386-5702	Sequence 5702, Ap
21	118	7.0	523	13	US-10-029-386-19512	Sequence 19512, A
22	109.8	6.6	523	13	US-10-029-386-8783	Sequence 8783, Ap
23	108.2	6.5	239	13	US-10-029-386-22509	Sequence 22509, A
24	93	5.6	2307	11	US-09-893-519A-87	Sequence 87, Appl
25	84.8	5.1	2886	15	US-10-156-761-1966	Sequence 1966, Ap
26	84.8	5.1	9025608	15	US-10-156-761-1	Sequence 1, Appl
27	83.4	5.0	3957	13	US-10-200-562-193	Sequence 193, App
28	83.4	5.0	3957	13	US-10-237-551-193	Sequence 193, App
29	83.4	5.0	154746	13	US-09-827-688-8	Sequence 8, Appl
30	83.4	5.0	154746	13	US-09-827-688-8	Sequence 8, Appl
31	81.8	4.9	125746	15	US-10-156-761-15102	Sequence 15102, A
32	81.8	4.9	9025608	15	US-10-156-761-1	Sequence 1, Appl
33	79.4	4.7	1392	13	US-10-214-446-7	Sequence 7, Appl
34	79.4	4.7	4689	10	US-09-861-289-34	Sequence 34, Appl
35	79.4	4.7	4689	10	US-09-860-846-34	Sequence 34, Appl
36	79.4	4.7	4689	11	US-09-988-384B-34	Sequence 34, Appl
37	79.4	4.7	4689	11	US-09-988-384B-34	Sequence 34, Appl
38	79.4	4.7	4689	13	US-09-836-821-34	Sequence 34, Appl
39	79.4	4.7	4689	13	US-10-271-889-34	Sequence 34, Appl
40	79.4	4.7	36778	10	US-09-860-846-5	Sequence 5, Appl
41	79.4	4.7	36778	11	US-09-836-821-5	Sequence 5, Appl
42	79.4	4.7	36778	13	US-10-271-889-48	Sequence 48, Appl
43	79.4	4.7	37948	11	US-09-988-384B-5	Sequence 5, Appl
44	79.4	4.7	38506	11	US-09-793-708-19	Sequence 19, Appl
45	79.4	4.7	38506	13	US-10-201-365-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
Sequence 21, Application US/09808387
Patent No. US20020132293A1
GENERAL INFORMATION:
APPLICANT: Kala Palm
APPLICANT: Tonis Timmek
TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
FILE REFERENCE: CEMRES.001A
CURRENT APPLICATION NUMBER: US/09/808,387
CURRENT FILING DATE: 2001-03-14
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 1675
TYPE: DNA
ORGANISM: Homo sapien
US-09-808-387-21

Query Match 100.0%; Score 1675; DB 10; Length 1675;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1675; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGGCAACACGCTGACCCGACCTTCCAGACCCGAGCCGCGCGCTCTGGCC	60
DB	1	ATGGGCAACACGCTGACCCGACCTTCCAGACCCGAGCCGCGCGCGCTCTGGCC	60
QY	61	ACCGGCGCTGCTGGCGCCCGCGCGCGAGGACGCGCGCTCTGGCGCGCGCGC	120
DB	61	ACCGGCGCTGCTGGCGCCCGCGCGCGAGGACGCGCGCTCTGGCGCGCGCGC	120
QY	121	TTCACCGGACGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG	180
DB	121	TTCACCGGACGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG	180
QY	181	CGGCGCAACAGCTTGCAGATGGCGTCACTTACGACGAGGCGCCATCCGGCTATACAG	240

```

Db      ||||| 181 CGGCGCAACAGCTTTCGCAATGGCGTCAAGTTCAAGGAGGCGCCATCCGGCTGTACGAG 240
Qy      ||||| 241 CAGGTGCGGCTGGGCTGGTGGCGGTGGCGGCTGGGCTGGAGCGGCGCTGGCTTCCGCG 300
Db      ||||| 241 CAGGTGCGGCTGGGCTGGTGGCGGTGGCGGCTGGGCTGGAGCGGCGGCTGGCTTCCGCG 300
Qy      ||||| 301 TTCACGCGGCAAGATCCGTCATGAGGCGCCAGAGCAATCCCAAGTAAGCGCTGCGCG 360
Db      ||||| 301 TTCACGCGGCAAGATCCGTCATGAGGCGCCAGAGCAATCCCAAGTAAGCGCTGCGCG 360
Qy      ||||| 361 GACCTGTATCAGCGGCGCGGCTTACTGGGCGAAGGCACTGCGCGAAGAACTGGGCTGCGC 420
Db      ||||| 361 GACCTGTATCAGCGGCGCGGCTTACTGGGCGAAGGCACTGCGCGAAGAACTGGGCTGCGC 420
Qy      ||||| 421 GACAGGTGTGGCTTACTGAGGCGGACCGGCAAGGCGCGGTTCATCAGCTGTGAACGAC 480
Db      ||||| 421 GACAGGTGTGGCTTACTGAGGCGGACCGGCAAGGCGCGGTTCATCAGCTGTGAACGAC 480
Qy      ||||| 481 GGGGAGCGGCTGCTTCCACTGCGGCGTGGCGGCGGCGGCGGCTGGGCGCTCATTT 540
Db      ||||| 481 GGGGAGCGGCTGCTTCCACTGCGGCGTGGCGGCGGCGGCGGCGGCTGGGCGCTCATTT 540
Qy      ||||| 541 GATGTCTACGAGCATCAACGAGGTGCAAGCTTCTGAGAGCGGCTTCTGCTGACAGCTG 600
Db      ||||| 541 GATGTCTACGAGCATCAACGAGGTGCAAGCTTCTGAGAGCGGCTTCTGCTGACAGCTG 600
Qy      ||||| 601 AGCGCGCGGCTGCTGAGCGGCGGCTTCAAGGCGCTGCTGCGGCGGCGGCGGCGGCGGCG 660
Db      ||||| 601 AGCGCGCGGCTGCTGAGCGGCGGCTTCAAGGCGCTGCTGCGGCGGCGGCGGCGGCGGCG 660
Qy      ||||| 661 GCGGCAACTTTCGACAAACAGAGCTGAGAACAAACAGAGTGTGGGCGAAGCTGGGCGAC 720
Db      ||||| 661 GCGGCAACTTTCGACAAACAGAGCTGAGAACAAACAGAGTGTGGGCGAAGCTGGGCGAC 720
Qy      ||||| 721 CTGGCGCTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
Db      ||||| 721 CTGGCGCTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
Qy      ||||| 781 TGGGGCGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Db      ||||| 781 TGGGGCGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Qy      ||||| 841 CGCTTTCACGCAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 900
Db      ||||| 841 CGCTTTCACGCAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 900
Qy      ||||| 901 GCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 960
Db      ||||| 901 GCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 960
Qy      ||||| 961 GAAAGCTCTTCTGAGAGGTGGGCGGCTGCGGCGGCTGGGCGGCGGCGGCGGCGGCGGCTTC 1020
Db      ||||| 961 GAAAGCTCTTCTGAGAGGTGGGCGGCTGCGGCGGCTGGGCGGCGGCGGCGGCGGCGGCTTC 1020
Qy      ||||| 1021 GGCATCAAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Db      ||||| 1021 GGCATCAAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Qy      ||||| 1081 GCGCTGTGCAAGCGCAAGAGTACTGGGTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140
Db      ||||| 1081 GCGCTGTGCAAGCGCAAGAGTACTGGGTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140
Qy      ||||| 1141 GGGGAGCGGCTGAGCTTCAAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
Db      ||||| 1141 GGGGAGCGGCTGAGCTTCAAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
Qy      ||||| 1201 CGTCCGCGGCGGCGGCTGCTGTGGGTGACACACAGAGGCGGCTTCTTCCGCG 1260
Db      ||||| 1201 CGTCCGCGGCGGCGGCTGCTGTGGGTGACACACAGAGGCGGCTTCTTCCGCG 1260
Qy      ||||| 1261 GTGCGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1320

```

```

Db      ||||| 1261 GTGCGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
Qy      ||||| 1321 ACCAGGACTCATGAGGAGTCCCTCAGCGGCTCCAGAGAGATGATGATTAATGAGAC 1380
Db      ||||| 1321 ACCAGGACTCATGAGGAGTCCCTCAGCGGCTCCAGAGAGATGATGATTAATGAGAC 1380
Qy      ||||| 1381 TTCAAGTGTACCAAGTCTCTCTCGAGCATCTGAGTCACTCTGAGTGAACGCGCCGAGCTTC 1440
Db      ||||| 1381 TTCAAGTGTACCAAGTCTCTCTCGAGCATCTGAGTCACTCTGAGTGAACGCGCCGAGCTTC 1440
Qy      ||||| 1441 CGGCTGAGGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 1500
Db      ||||| 1441 CGGCTGAGGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 1500
Qy      ||||| 1501 GCGGAGTGAACGAGTGTCTTCATGAGGAGTGAACAGGATCACTACAGTGTGAGAC 1560
Db      ||||| 1501 GCGGAGTGAACGAGTGTCTTCATGAGGAGTGAACAGGATCACTACAGTGTGAGAC 1560
Qy      ||||| 1561 ATGTGCTGTGCAACAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1620
Db      ||||| 1561 ATGTGCTGTGCAACAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1620
Qy      ||||| 1621 ATCTGCGGCGGCGGCGGCTGCGGCGGCGGCGGCTGCGGCGGCGGCGGCTGCGGCGGCGG 1675
Db      ||||| 1621 ATCTGCGGCGGCGGCGGCTGCGGCGGCGGCGGCTGCGGCGGCGGCGGCTGCGGCGGCGG 1675

RESULT 2
US-09-808-387-27
; Sequence 27, Application US/09808387
; Patent No. US2002013293A1
; GENERAL INFORMATION:
; APPLICANT: Kala Palm
; APPLICANT: Tonle Tjmmuak
; APPLICANT: Cemines Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; TITLE OF INVENTION: TRANSCRIPTION REGULATORS AND USES THEREFOR
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808,387
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1641
; TYPE: DNA
; ORGANISM: Rat
US-09-808-387-27

Query Match      73.3%; Score 1228; DB 10; Length 1641,
Best Local Similarity 84.3%; Pred. No. 8, 8e-263;
Matches 1406; Conservative 0; Mismatches 235; Indels 27; Gaps 1;

Qy      ||||| 1 ATGGGCAACAGGTGACCGGACCGTCCGACAGCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 60
Db      ||||| 1 ATGGGCAACAGGTGACCGGACCGTCCGACAGCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 60
Qy      ||||| 61 ACCCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 120
Db      ||||| 61 ACCCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 120
Qy      ||||| 121 TTCACGCGCAAGCGCAAGAGGAGAAAGTGGCGGCTGAGACGCGCACTTGGCGGCGGCGCA 180
Db      ||||| 121 TTCACGCGCAAGCGCAAGAGGAGAAAGTGGCGGCTGAGACGCGCACTTGGCGGCGGCGCA 180
Qy      ||||| 181 CCGCGCAACAGCTTCTGCAATGGCGTCAAGTTCAGCAAGCGGCGGCGGCGGCGGCGGCGGCG 240
Db      ||||| 181 CCGCGCAACAGCTTCTGCAATGGCGTCAAGTTCAGCAAGCGGCGGCGGCGGCGGCGGCGGCG 240
Qy      ||||| 241 CAGGTGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
Db      ||||| 241 CAGGTGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
Qy      ||||| 301 TTCACGCGCAAGATCCGTCATGAGGCGCCAGAGCAATCCCAAGTAAGCGCTGCGCG 360

```

Db 301 TTCACTGCAAGACCGGTGCTCATAGTCGACAGATATCCCAAGTACGCTGCCCC 360
Qy 361 GACCTGGTCAAGCGGCGGCTACTGAGGCAAGGCACTGCGGAGAACTGGCGTGGC 420
Db 361 GACCTGGTCAAGCGGCGGCTACTGAGGCAAGGCACTGCGGAGAACTGGCGTGGC 420
Qy 421 GACAGGCTGCTGAGCTTACTGAGGCGGACCGGCGGCGGCTTCTACAGGCTGAGAC 480
Db 421 GACAGGCTGCTGAGCTTACTGAGGCGGACCGGCGGCGGCTTCTACAGGCTGAGAC 480
Qy 481 GAGCAGCGGCTGCTTCTCATCTGAGGCGTGGCCGTGGCGGCGGCTTGGCGCTCAT 540
Db 481 GAGCAGCGGCTGCTTCTCATCTGAGGCGTGGCCGTGGCGGCGGCTTGGCGCTCAT 540
Qy 541 GATGCTTACGCGCATCACCGAGAGGTGAGGCTTCTGAGAGCGGCTTCTGCTGACAGCTG 600
Db 541 GATGCTTACGCGCATCACCGAGAGGTGAGGCTTCTGAGAGCGGCTTCTGCTGACAGCTG 600
Qy 601 AGCGCGCGCGCTTCAAGCAGGCGCGCTTCAAGCGCTTCAAGCGCGCGCGCGCGCGCG 660
Db 601 AGCGCGCGCGCTTCAAGCAGGCGCGCTTCAAGCGCTTCAAGCGCGCGCGCGCGCGCG 660
Qy 661 GCGGCGCACTTGCACACAGAGCTTCAAGAGCAACAGGCTGAGGCGCACTGAGGCGCAC 720
Db 661 GCGGCGCACTTGCACACAGAGCTTCAAGAGCAACAGGCTGAGGCGCACTGAGGCGCAC 720
Qy 721 CTGGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
Db 721 CTGGCGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
Qy 781 TGCGGCGCGCGCTTCAAGCAGGCGCGCGCTTCAAGCGCGCGCTTCAAGCAGGCGCG 840
Db 781 TGCGGCGCGCGCTTCAAGCAGGCGCGCGCTTCAAGCGCGCGCTTCAAGCAGGCGCG 840
Qy 841 CGCTTCCACGCAACAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 841 CGCTTCCACGCAACAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy 901 GCACCG 960
Db 901 GCACCG 960
Qy 961 GAGAGCGCTTCTGAGAGGTGAGGCGCGCTTCAAGCGCGCGCGCGCGCGCGCGCG 1020
Db 961 GAGAGCGCTTCTGAGAGGTGAGGCGCGCTTCAAGCGCGCGCGCGCGCGCGCGCG 1020
Qy 994 GGCATCAAGCTGCTGAGATCTGAGCGCGCTGCGCGCATCCGAGCTTCCGCGCT 1053
Db 994 GGCATCAAGCTGCTGAGATCTGAGCGCGCTGCGCGCATCCGAGCTTCCGCGCT 1053
Qy 1081 GCGGCTGCTGAGCG 1140
Db 1081 GCGGCTGCTGAGCG 1140
Qy 1141 GCGGCGCGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 1200
Db 1141 GCGGCGCGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 1200
Qy 1201 GCGGCG 1260
Db 1201 GCGGCG 1260
Qy 1261 GCGGCG 1320
Db 1261 GCGGCG 1320
Qy 1321 ACACGAGCTTCAAGGCTTCTTCAAGGCTTCTTCAAGGCTTCTTCAAGGCTTCTTCA 1380
Db 1321 ACACGAGCTTCAAGGCTTCTTCAAGGCTTCTTCAAGGCTTCTTCAAGGCTTCTTCA 1380
Qy 1381 TTCACTGCAAGACCGGTGCTCATAGTCGACAGATATCCCAAGTACGCTGCCCC 1440
Db 1381 TTCACTGCAAGACCGGTGCTCATAGTCGACAGATATCCCAAGTACGCTGCCCC 1440

Db 1354 TTGCGGCTCAACAGTGTATCATGAGTCAAGATCTGCTGAGACAGGCGCGCGCTCC 1413
Qy 1441 CCGCTGAGCG 1500
Db 1414 CCACTGAGTGTGCG 1473
Qy 1501 GCGGAGTGCACAGGCTGCTTCAAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1560
Db 1474 GAGAGTGCACAGGCTGCTTCAAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1533
Qy 1561 ATGAGCTGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1620
Db 1534 ATGAGCTGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1593
Qy 1621 ATGAGCTGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1688
Db 1594 ATGAGCTGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC 1641

RESULT 3
US-09-808-387-23
Sequence 23, Application US/09808387
Patent No. US20020132293A1
GENERAL INFORMATION:
APPLICANT: Kala Palm
APPLICANT: Tom's Timmusk
TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
TITLE OF INVENTION: TRANSCRIPTION REGULATORS AND USES THEREFOR
FILE REFERENCE: CEMRS 001A
CURRENT APPLICATION NUMBER: US/09/808,387
CURRENT FILING DATE: 2001-03-14
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 1129
TYPE: DNA
ORGANISM: Homo sapien
US-09-808-387-23

Query Match 65.9%; Score 1103.6; DB 10; Length 1129;
Best Local Similarity 98.8%; Pred. No. 3,1e-235;
Matches 1112; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 550 GGCATCAAGCTGCTGAGATCTGAGCGCGCTTCAAGCGCGCGCGCGCGCGCGCG 609
Db 4 GGCATCAAGCTGCTGAGATCTGAGCGCGCTTCAAGCGCGCGCGCGCGCGCGCG 63
Qy 610 GCGCTGAGCG 669
Db 64 GCGCTGAGCG 123
Qy 670 TTGACAAACAGAGCTTCAAGAGCAACAGAGTGCAGAGTGCAGAGTGCAGAGTGC 729
Db 124 TTGACAAACAGAGCTTCAAGAGCAACAGAGTGCAGAGTGCAGAGTGCAGAGTGC 183
Qy 730 GCG 789
Db 184 GCG 243
Qy 790 GCGGAGCG 849
Db 244 GCGGAGCG 303
Qy 850 GCAACAGCG 909
Db 304 GCAACAGCG 363
Qy 910 CCGGAGCG 969
Db 364 CCGGAGCG 423
Qy 970 TTGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAG 1029

```
Db 424 TTGTGAGAGTGAGGCGTCCGAGGCTGGCGGCGCGGCTGGGCTTGCGGCATACG 483
1030 TCGTGGAGACCCGGGCGGTGTAAGGCCCAAGAGCTGGCGGCGGCGGCTGGGCTGTC 1089
484 TCGTGGAGACCCGGGCGGTGTAAGGCCCAAGAGCTGGCGGCGGCGGCGGCTGGGCTGTC 543
1090 GACCGCAAGAGTACTGGGTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1149
544 GACCGCAAGAGTACTGGGTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 603
1150 CTCAGCTTCACTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1209
604 CTCAGCTTCACTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663
1210 GGGCGCTGCTGGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1269
664 GGGCGCTGCTGGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723
1270 GGGCGCTGCTGGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1329
724 GGGCGCTGCTGGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 783
1330 CCATCAGAGTCTCTCAGCGGCTCCAGAGCATATGATTCAGATGATGATGATGATGATGATG 1389
784 CCATCAGAGTCTCTCAGCGGCTCCAGAGCATATGATTCAGATGATGATGATGATGATGATG 843
1390 AACAGTCTCTCTGAGTCAATCCCTGGTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1449
844 AACAGTCTCTCTGAGTCAATCCCTGGTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 903
1450 CCCCCGCTGCTCCCGGCTCTCCCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1509
904 CCCCCGCTGCTCCCGGCTCTCCCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 963
1510 ACCGTGTCTTCATGAGTGGAGTGAACAGGTATCTTACAGTGTGAGCAATGTGCTG 1569
964 ACCGTGTCTTCATGAGTGGAGTGAACAGGTATCTTACAGTGTGAGCAATGTGCTG 1023
1570 TGCACAGCTGGCGGCTGGCGGCTGAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1629
1024 TGCACAGCTGGCGGCTGGCGGCTGAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1083
1630 CGGCGCATCAAGAGCTCATTAAGATCTAAGGCGCATAGCTAGCC 1675
1084 CGGCGCATCAAGAGCTCATTAAGATCTAAGGCGCATAGCTAGCC 1129

RESULT 4
US-09-808-387-25
; Sequence 25, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemines Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808.387
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-808-387-25

Query Match 34.5%; Score 577.2; DB 10; Length 955;
Best Local Similarity 99.5%; Pred. No. 9.9e-119;
Matches 579; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 ATGGGCAACAGGTGACCGGACCTTGCCAGAACCCGAGCCCAACCGGCGGCTCTGCGCC 60
Db 1 ATGGGCAACAGGTGACCGGACCTTGCCAGAACCCGAGCCCAACCGGCGGCTCTGCGCC 60
QY 61 ACCCGGCGGCTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 ACCCGGCGGCTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 TTCACAGCGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 180
Db 121 TTCACAGCGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 180
QY 181 CCGCGCAAGAGCTTCTGCAATGAGCGTCAAGGCTTCAAGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 CCGCGCAAGAGCTTCTGCAATGAGCGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 241 CAGGTGGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 241 CAGGTGGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 301 TTCACCGCGCAAGATCCGTGCTCATAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
Db 301 TTCACCGCGCAAGATCCGTGCTCATAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 361 GACCTGTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db 361 GACCTGTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 421 GACAGGAGTCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Db 421 GACAGGAGTCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 481 GGGAGCGGAGTCTTTCATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 540
Db 481 GGGAGCGGAGTCTTTCATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 540
QY 541 GATGTCTACGGCATCACCGAGAGTGCAGCTTCTGAGAGC 582
Db 541 GATGTCTACGGCATCACCGAGAGTGCAGCTTCTGAGAGC 582

RESULT 5
US-09-808-387-3
; Sequence 3, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemines Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808.387
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-808-387-3

Query Match 24.4%; Score 409.2; DB 10; Length 1674;
Best Local Similarity 56.6%; Pred. No. 1.4e-81;
Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;
```

```
QY 99 GGTCTGGGCGGAGGCGGCGGCTTTCACGCGGAGGCGGCAAGGCAAGGCAAGGCAAGGCAAG 158
Db 120 GCTCCAGGCAACCGGCTGCTTTCACCGGCAACGAGGCTCCCAATCTCATGGA 179
QY 159 CGGCACTGCGCGGCGGCGGCAACGCGGCAACGAGCTTCTGCAATGAGGCTCATGTCGCA 218
Db 180 CTTGAGCCACAGAGGCTGTCAAGAGGAGGCGGAGGCTTCTGCAAGCCATCACTTCAAGCA 239
```

QY 219 GCGGCCATCGGCTGTACAGAGAGGTGCGGCTGCGCTGTGCGCCCTGCGCTG 278
 DB 240 CCGCCCGGCTCTCATCTACAGAGAGGTGCGGCTGAGATCAACAGAGAGGTGCGCTG 299
 QY 279 GAGGCGGCGGCTGCGCTGTGCGCTGTACCGGCGAGATCGGCTGTGAGCGGCGCA 338
 DB 300 GAGGCGGCGGCTGCGGCTGTGCGCTGTACAGAGAGAGGTGCGGCTGTGAGCTC 359
 QY 339 CATCCCAAGTACCGCTGCGGCGAGCTGTGTACCGGCGGCGGCTGTGAGCGGCGCA 398
 DB 360 GGTGCGCAAGTACCGCTGCGGCGAGCTGTGTGTCCAGAGGTGCGCTGTGAGCGGCGCT 419
 QY 399 GCGCGAGAACCTGCGGCGTGTGCGGCGAGAGGTGTGCGCTGTGCGGCGAGCGGCGG 458
 DB 420 GCGTGAAGAGTGTGCGCAATGAGGCGCAATCATGCGATTCGTGCGTGTGCAAGAGGCGG 479
 QY 459 CGTGTGTACAGCGGTGAGAGCGGCGGCGGCGGCTGTGTCCAGCGGCGGCGGCGG 518
 DB 480 TGTCTTCAACCGCATCAAGAGCTGCGCTGTGTATGCTGTTCAGCGGCGGCTGTGAGCGG 539
 QY 519 GCGCGCGCTGTGCGGCTGTGAGTGTGTACAGGATCAACGAGAGGTGAGCTTGTGAG 578
 DB 540 GAGCGGCTGTGCGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
 QY 579 GAGGCGCTGT 638
 DB 600 TAGGAGCTGTG-----TGTCTCCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 650
 QY 639 CTGTGCGGCGGCGAGAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 698
 DB 651 GCGCGGCGGCGGCTGT 699
 QY 699 GGT 758
 DB 700 -----AGCTTATGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 749
 QY 759 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 818
 DB 750 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 809
 QY 819 GCGCGT 875
 DB 810 GCGCGAGCTGT 869
 QY 876 GTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 935
 DB 870 CTGT 929
 QY 936 CTGT 995
 DB 930 CACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 989
 QY 996 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1055
 DB 990 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1049
 QY 1056 CAAGAGCTGT 1115
 DB 1050 GCGCGAGCTGT 1109
 QY 1116 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1175
 DB 1110 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1169
 QY 1176 GAGAGT 1235
 DB 1170 GAGAGT 1229
 QY 1236 GAGAGGCTGT 1295
 DB 1230 GAGAGGCTGT 1289

QY 1296 GAGTACCTGT 1346
 DB 1290 CTGT 1349
 QY 1347 -----GCGTCCAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1397
 DB 1350 AACGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1409
 QY 1398 CTGT 1457
 DB 1410 TGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1469
 QY 1458 GTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1517
 DB 1470 GTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1520
 QY 1518 CTGT 1577
 DB 1521 CTGT 1580
 QY 1578 CTGT 1637
 DB 1581 CTGT 1640
 QY 1638 CAAGAGCTGT 1661
 DB 1641 CAAGAGCTGT 1664

RESULT 6

US-09-808-387-1
 ; Sequence 1, Application US/09808387
 ; Patent No. US20020132293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kaia Palm
 ; APPLICANT: Tonis Timmusk
 ; APPLICANT: Cemines Research
 ; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
 ; FILE REFERENCE: CEMRES.001A
 ; CURRENT APPLICATION NUMBER: US/09/808,387
 ; CURRENT FILING DATE: 2001-03-14
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PatsSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1725
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-808-387-1

Query Match 24.4%, Score 409.2; DB 10; Length 1725;
 Best Local Similarity 56.6%; Pred. No. 1.4e-81;
 Matches 896; Conservative 0; Mismatches 628; Indels 60; Gaps 5;

QY 99 GGT 158
 DB 171 GGT 230
 QY 159 GCGGCACTGT 218
 DB 231 CTGT 290
 QY 219 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 278
 DB 291 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 350
 QY 279 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 338
 DB 351 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 410
 QY 339 CATCCCAAGTACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 398
 DB 411 GGTGCGCAAGTACCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 470

QY 399 GCCCGAAGAACTGACGCTGCGGACACAGGTGCTGAGCTTACTGAGCCGACCGCCG 458
 DB 471 GCTTGAAGAGTTTGGCCATGAGGGGCAATCATTCGATTTCTGGGTGGACAAGAGGCGG 530
 QY 459 CGTGTCTTCAAGCGTGAACGACGCGGAGCGGCTCTTTCACCTGCGGCGTGGCCGTGG 518
 DB 531 TGTCTTCCACCGCATCAACGACTCGGCTGTATGCTGTTCTTCAGCGGGGTCCGACGCG 590
 QY 519 CGGCGGCTCTGAGGCGCTATTTATGTCTAACCGGATCAACGAGAGGTGACGTTTGA 578
 DB 591 CACCGGCTCTGAGGCGCTGTGAGCGTCTAACCGGCTCACGCGGGGCTCAGGCTGTGA 650
 QY 579 GAGCGGCTTCCGCTGACACGCTGACGCGCGCGCTCAGCGAGGCGCGCTTACGCGCTG 638
 DB 651 TAGCGAGCTGG-----TGTCTCCGAACTGTCTGCGGCGCGGCTCTTCAACCGCCCT 701
 QY 639 CTTGCGGCGCCAGACGACGACGCGGCGCACTTTCGACAAACAGAGCTGAGAGAACCA 698
 DB 702 GCGGCGGCGCGTCCGCTGCGGCGGAGAGCGGACGACGCGCGCTCTCGGTG----- 750
 QY 699 GGTGTGCGCAAGCTGAGGCGCACTTGGCGCTGAGGCGCGCGCGCGCGCGCGCGCG 758
 DB 751 -----AGCTATGCGACTTCAACGTCGCGGCGCGGACGCGGCGCGCGCGCG 800
 QY 759 CGCGCGGCGCGCGCGCATTCCTGTGCGGCGCGCGTGAAGCGCGCGCGCGCGCTGTGCG 818
 DB 801 GCGCGCGGCTGCGCGCATCCCGAGAACTCACTCACTGCGGACAGCGCGCGCGCTGCG 860
 QY 819 GCGGCTACTGAGAGCGGACCTGCGCTTTCACGCGAACAGCGCGGCGCG---ACGTGAGCT 875
 DB 861 GCGGCGAGCTGACGCGGACCTGCGGTTTCCAGCGCTGCGGCGCGGCGCGCGACGTGCG 920
 QY 876 GTGCGGCGGACCGGAAAGTGGCTTGGCGGACCGCGCGCGCGCGCGCGCGCGCGCG 935
 DB 921 CTTGACGAGACAGACGCGTGGCGCGCTGTGAGGACGCGGCGCGGACGCGCTCGCTT 980
 QY 936 CTTGAGAGCGCGCGCTGCGCGCGCGCGCGGAGAGCTCTTCTGTGAGAGTGGCGCGT 995
 DB 981 CACCAAGCGCGCGCGCTGCGCGCGCGCGCGGACGACGCTTCTGTCAAGTCAAGCGCT 1040
 QY 996 GCGGCGCGCGCGCGCGCTTTCGACATCAGCTGTGAGGACCGCGGCGGTCTACGCG 1055
 DB 1041 CGGCGCGCGCGCGCGCTGCTGTGCGCGCTCACAGGTGCGACCGCGCGCGCGCTGCG 1100
 QY 1056 CAACGAGCTGCGCGCGCGGACCGGACGCGCTGTGACCGGAAAGAGTCTGGGTGAG 1115
 DB 1101 GCGCGACCTGCTTTTACGCGCTGTGAGCGCTGTGAGCGGAAATTTCTGGCGCG 1160
 QY 1116 GCGCGCGCGCGCGCGCTGCGGAGCGCGCGGACGCGCTCAGCTTCAAGCTGCGCGCG 1175
 DB 1161 CGCGGTGCGCGCGCGCGCTGTGACAGCGCGGACGATCTGGGCTGTGTCAACCGCA 1220
 QY 1176 CAGAGTCTCTGAGGACATCAACGCGCGTCCGCGCGCGCGCTGTGTGTGACACAC 1235
 DB 1221 CGAGCTGACGCTCAGGCAACATGCGCGCGCGCGCGGACATGACGTGTGACGCGCT 1280
 QY 1236 GAGAGGCTCTGAGGCGCTTCTTCCCGCGGCGCGCGCGCGGTGAGCGGCGAGTGTGCT 1295
 DB 1281 GAGCGCGCTTGTGAGTCTTTCGCGCTGTGACGCGGACATGACGAGTCTCGCAT 1340
 QY 1296 CGGTACCTGTGACGCTGACGCGCTGACGACGAGTCTTCAAGGTCTCTCAG----- 1346
 DB 1341 CTCACATATCTTGGCGGAGCGGAGTATCCGTGATCTCCCTGCTCCCTGCGACGCG 1400
 QY 1347 -----CGGCTCCGAGAGCATAGTATGATTTCAATATGACCTTCACTGTCAACCA 1397
 DB 1401 AACCTGCGGAGGCGCTGAGGAGCGCGCTGTGACCGCTTGTCTGACGAGCTC 1460
 QY 1398 CTCTGCGGACATGAGTCACTCCGTGAGGAGCGCGCGCGCGCTCCGCTGAGCGCGCG 1457
 DB 1461 TGGCGCTTGTGAGTCTTGTGTGTGTGAGGAGCGCGCGCGCAATTCGCAATGAGCTG 1520

QY 1458 GTCCCCCGTGTCTTCCCAACCGGAGCGCGGACGATCAAGATGCGAGTGCAGGTG 1517
 DB 1521 GTGCGCAATGACCCCAAGGTCTGAGGCAATG-----GAGCGATGAGTGACCATTTG 1571
 QY 1518 CTTGATGCGGAGGTGAGCAGGTGATCTTACACGATGAGGACATATGCTGTGCGACAG 1577
 DB 1572 CTATGAAACGCGGTGAGACCGGTATCTTACATATGAGGACATATGCTGTGCGAC 1631
 QY 1578 CTGCGGCTGCGGCTCAAGCGGACGAGCGCGGCTGTGCGCGCATCTGCGCGCGCG 1637
 DB 1632 CTGTGCGCTGCGGCTCAAGAGGCTGTGACGCGCTGTGCGCGCATCTGCGCGCG 1691
 QY 1638 CAAGAGCTGATTAAGTCTACAG 1661
 DB 1692 CAAGAGCATCATCAAGACTACG 1715

 RESULT 7
 US-09-808-387-9
 ; Sequence 9, Application US/09808387
 ; Patent No. US20020132293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kala Palm
 ; APPLICANT: Tonis Timmusk
 ; APPLICANT: Cemines Research
 ; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
 ; FILE REFERENCE: TRANSRIPTION REGULATORS AND USHS THEREFOR
 ; CURRENT APPLICATION NUMBER: US/09/808,387
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: FastSB for Windows Version 4.0
 ; SEQ ID NO 9
 ; TYPE: DNA
 ; ORGANISM: mouse
 US-09-808-387-9

 Query Match 23.4%; Score 391.6; DB 10; Length 1698;
 Best Local Similarity 55.9%; Pred. No. 1.1e-77;
 Matches 885; Conservative 0; Mismatches 639; Indels 60; Gaps 5;

 QY 99 GGTCTGTGAGGCGCGCGCTTTCACGCGGCGGCAAGGCAAGAGCTGCGGCTG 158
 DB 120 GCTCCCGGACGCGCGCTGCTTCCACCGCGCACTAAGGCGCTCCAGATCTCATG 179
 QY 159 CGGCACTGCGCGCGCGCGGCAACGCGGCAAGCTTCTGCAATGCGTCAAGCTCA 218
 DB 180 CTTGAGGCAAGGCGCGCTGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 239
 QY 219 GCGGCGCATTCGCGCTGTGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 278
 DB 240 CGCGCGGCGCTCATCTTACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 299
 QY 279 GAGCGGCGGCGGCGGCTTGTGCTTACCGGCGGCAAGTCCGTGTGCTTACGAGCG 338
 DB 300 GAGCGGCGGCGGCGGCTTGTGCTTACCGGCGGCAAGTCCGTGTGCTTACGAGCG 359
 QY 339 CATCCCAAGTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 398
 DB 360 GCTGCGCAAGTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 419
 QY 399 GCCCGAAGAACTTGGCGCTGCGGCGGACAGGTGTGCTTACGAGCGGCGGCGGCG 458
 DB 420 GCTGAGAGAGTTTGGCAAGAGGCGGCAATCATTTGCTTCTGAGGTGACAAAGAGG 479
 QY 459 CGTGTCTTCAAGCGTGAACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 518
 DB 480 CGTCTTCTACCGGATCAAGTACGCTGTGCTTCTTCTTCAAGTGGGAGTCCGAGCG 539
 QY 519 CGGCGCGCTGCGGCGGCTCATTTGATGTCTACGCGGATCAAGGAGGTGACGCTTGT 578
 DB 540 GAGCGGCTCTGCGGCGGCTGTGTGAGCGTCTACGCGGCTCACGCGGAGTGTCACTG 599


```

QY 759 GCGCGCGCGCGCGCATTCGGTGGGGGCGCGGTGAGGCGCGCGCGCGCGTCCGCC 818
DB 801 GCGTCCCGGCGTGGCGCGATCCCGGAGAACTGGCTCAATTCTTCAAGACAGCGCGCGTGGC 860
QY 819 GCGCGTACTGAGAGCGGACCTGCGCTTTCACGCAACAGCGGCGCGCG---ACGTGAGCCT 875
DB 861 GCGCGAGCTCGACGCGGAGCTGCGCTTTCACGCGCTTTCGCGCGCGCGCGCGAGTCCGCA 920
QY 876 GTGCGCGGACCGGAAAGTGGCTGCGCACCGCGCGCGCGCGCGCGCGCGCGTCTT 935
DB 921 CTTGAGACGAGACAGCGTGGCGCGCTTGGAGACAGCGCGCGCGCGCGCGCGCTT 980
QY 936 CTTCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 995
DB 981 CACGAGCGCGCGCTTGGAGCGTGGCGCGAGACATCTTCAATCAAGGTACGCGCGCGCG 1040
QY 996 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1055
DB 1041 GCGCGGAGCGCGCGCGCGCGCTTCTTGGGGGTCAACAGGTGAGCCGCGCGCGTGGCG 1100
QY 1056 CAACGAGCTCCCGCGCGCGCGCGCGCGCGCTGCGACCGGAAAGATCTGGGTGGC 1115
DB 1101 CCGCGGACCTGCGCTTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTG 1160
QY 1116 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1175
DB 1161 TGGCGGCGCGCGCGCGCTTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1220
QY 1176 CGAGGTGCTCTGGGCGATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1235
DB 1221 AAGAGCTGACCTGAGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1280
QY 1236 GCGGCGCTGTTGGCGCTTCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1295
DB 1281 GCGGCGCTGTTGGCGCTTCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1340
QY 1296 CGGTACCGTGGAGTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1342
DB 1341 CTTCACCATCATGATGAAAGCGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCG 1400
QY 1343 -----TCAGCGCGCTCCAGGAGCATATGATTCAGATATGACCTTCAAGTCAACGATC 1397
DB 1401 AACCTCAACGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1460
QY 1398 CTCTCGGCGATCTGATCATCTCTGGTGAAGCGCGCGCGCGCGCGCGCGCGCG 1457
DB 1461 TGGGCGCGCTGAGTGGCTGCTGAGAGGAGACGCGCGCGCGCGCGCGCGCG 1520
QY 1458 GTCCCGCGGTCTCTCCCAACCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 1517
DB 1521 GTCAACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1571
QY 1518 CTTGATGAGCGAGTGGAGACAGGTCACTTACAGTGGAGACAGTGGCGCGCG 1577
DB 1572 CTATGAAACAGGAGTGGATGATGATCAAGTGGAGACAGTGGCGCGCGCG 1631
QY 1578 CTGCGGCGTGGCGCTCAAGCGAGACGCGCGCGCGCGCGCGCGCGCGCG 1637
DB 1632 CTGTGCGCTGGCGCTCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1691
QY 1638 CAAGGAGCGTATTAAGATCTACAG 1661
DB 1692 CAAGGAGCATCATCAAGACTTACCG 1715

```

```

RESULT 9
US-09-808-387-15
; Sequence 15, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kala Palm
; APPLICANT: Tonis Timmusk

```

```

/ APPLICANT: Cemine Research
/ TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
/ TITLE OF INVENTION: TRANSCRIPTION REGULATORS AND USES THEREFOR
/ FILE REFERENCE: CERES. 001A
/ CURRENT APPLICATION NUMBER: US/09/808.387
/ CURRENT FILING DATE: 2001-03-14
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15
/ LENGTH: 1725
/ TYPE: DNA
/ ORGANISM: rat
US-09-808-387-15

Query Match 22.7%; Score 380.8; DB 10; Length 1725;
Best Local Similarity 54.7%; Pred.No. 2.7e-75;
Matches 903; Conservative 0; Mismatches 687; Indels 60; Gaps 5;

QY 33 CCGGAGCGCGCGCGCGCGCGCTTCTGGCGACCGCGCGCGCGCGCGCGCGCGCG 92
DB 105 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 164
QY 93 AGCGCGGTCTGGGCGAGGCGCGCGCGCTTCAAGCGCGCGCGCGCGCGCGCG 152
DB 165 TGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 224
QY 153 GCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 212
DB 225 CATGAGACCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 284
QY 213 CAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 272
DB 285 CAGGAAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 344
QY 273 TGGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 332
DB 345 CTGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 404
QY 333 CGAGGACATCCCGCGAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 392
DB 405 GCGACTACCTCCCGAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 464
QY 393 GCGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 452
DB 465 AGCGTGGCTGAGAGGTTTGGCGAGGCGCGCGCGCGCGCGCGCGCGCG 524
QY 453 CCGCGCGGTGTTCTACAGCGTGAAGCGCGCGCGCGCGCGCGCGCGCG 512
DB 525 GCGCGGAGTCTTCAACCGGATCAAGAGTGGCGCGCGCGCGCGCGCGCG 584
QY 513 CGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 572
DB 585 AACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 644
QY 573 TCTGAGAGCGCGCTTCTGCTGACAGCGCGCGCGCGCGCGCGCGCGCG 632
DB 645 GCTTAAAGCG-----GAGCTGGTGGCTGACAGCGCGCGCGCGCGCG 695
QY 633 CCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 692
DB 696 CCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 742
QY 693 CAACGAGTGGGCGCAAGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 752
DB 743 -----TGTCTGAGAGCTTGGGAGCTTCAAGTGGCGGAGCGCGGAGGAG 794
QY 753 AGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 812
DB 795 GCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 854
QY 813 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 869
DB 855 GCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 914

```

```

Qy      870 GAGCTGTGCGCCGACCCGAAAGTGGCTTGCGCACCGCGGCCCGACCGCGCGCCCT 929
Db      915 CCGCATCTGTGACGAGCAGACGGTGGCGCGCTTAAGACACGGGCGCGACGAGCGCGCCT 974
Qy      930 GATCTTCTCCGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 989
Db      975 GGTCTTTCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1034
Qy      990 GGGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1049
Db      1035 AGGCGGTGCGCGACGAGAAAGCGGTGCTCTTGGGGGTGACACAGTGTGACCGTGGACGCT 1094
Qy      1050 ACGGCGCCAAACGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1109
Db      1095 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1154
Qy      1110 GATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1169
Db      1155 GGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1214
Qy      1170 CGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1229
Db      1215 GACGCGAAAGCTGACCTTAATTACAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1274
Qy      1230 CACCAACGAGCGCGCTGTGGCGCTTTCTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1289
Db      1275 GCGCTGCGACGCGCGCTGTGAATGCTTTCAAGCGCTGACCGCGCGCGCGCGCGCGCGCG 1334
Qy      1290 TCTCTCGGTACCGCTGACGTCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1346
Db      1335 CTTTGTGCTCAACATCATAGGTGAAGCGGGTGGCCCATCTCTCCCTGCTACCTGCGCTC 1394
Qy      1347 -----CGGCTCCACAGACGATAGTATTCAATATGACCTTCAAGTGA 1391
Db      1395 CACTCCAACTCGCGCAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1454
Qy      1392 CCAAGTCTCTCGCGCATGTAGTATCTCCGTGTGACGGCGCGCGCGCGCGCGCGCGCGCGCG 1451
Db      1455 CGATTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1514
Qy      1452 CCGCGTGTCCCGCGGTCTCTCCCAACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1511
Db      1515 GCCGAGTACCAAGTACCGCGCGGTG-----GACAGTGAACGATGAATGAC 1565
Qy      1512 GGTGTGCTTCAGTGGCGAGTGGACACCGGTATCTTACAGTGTGACACATGTGCTGTG 1571
Db      1566 CATTTGTATGAACAACGAGTGGATACATGATCATACATGTGTGCGACATGTGTGTG 1625
Qy      1572 CCAACAGTGGCGCGCTGTGGCGCTTCAAGCGACAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1631
Db      1626 CTACTCTGTGGCGCGCGCGCTTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1685
Qy      1632 GCCCATCAAGACGTATTAGATCTACG 1661
Db      1686 CCCCATCAAGACATCATCAAGACCTACCG 1715

```

```

RESULT 10
US-09-808-387-5
; Sequence 5, Application US/09808387
; Patent No. US2002013229A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemines Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: TRANSCRIPTION REGULATORS AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/808.387
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PastSeq for Windows Version 4.0

```

```

; SEQ ID NO 5
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-808-387-5

Query March 15.84; Score 264; DB 10; Length 1161;
Best Local Similarity 56.64; Pred. No. 2e-49;
Matches 567; Conservative 0; Mismatches 405; Indels 30; Gaps 3;

Qy      681 CAGACTCGAAGAACACAGGTGTGTGCGCAAGCTGGGCGACCTGGCGCTGAGCGCGCGCGCG 740
Db      159 CAGAGCGAGAACACCGCGCGCTCTGCGTGAAGCTATGGAACCTTCAACAGTGGCGCGCGCG 218
Qy      741 GGGGCCACCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 800
Db      219 CCGGACGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 278
Qy      801 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 860
Db      279 GCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 338
Qy      861 GCCCG---ACGTGAAGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 917
Db      339 CCGGCGCGACGTCGCGCATCTCGACGAGACAGACGAGTGGCGCGCGCGCGCGCGCGCGCGCG 398
Qy      918 CCGCGCGACGCTGTCTTTCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 977
Db      399 CGAGCGCGCGCTCGCTTTCACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 458
Qy      978 GGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1037
Db      459 GGTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 518
Qy      1038 CCGCGCGGTGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1097
Db      519 CCGCGCGACGCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 578
Qy      1098 AGAGTATGTGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1157
Db      579 GGAATTTCTGGGCGCGGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 638
Qy      1158 CACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1217
Db      639 GGTGTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 698
Qy      1218 GCTGTGCGTCAACACAGCGCGCGCTGTGGCGCTTCTTGTGCGCGCGCGCGCGCGCGCGCGCG 1277
Db      699 GCTGTGCGTGAACCGCTGTGACGCGCGCTTGTGAATGCTCTTGGCGCTGACCGGACCATAC 758
Qy      1278 GGGCGAGTGTGTCTCTGTGGTACCTGTGACGTCAAGCTTGTGACACAGACTCATACG 1337
Db      759 GCAATCGGCATCTCGGCTCACTATCTGGCGGAGGGGGGTATCCGTCATCTCCCTG 818
Qy      1338 GTCCCTCAGCGCGT-----CCCAAGAACATAGTATTAGATATGAC 1379
Db      819 CTCCCTGTGCTCAACGCGCAACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 878
Qy      1380 CTTGAGTGTCAACAGTCTCTCTCGGCACTGAGTATCTCTGTGTGAACGCGCGCGCGCGCGCG 1439
Db      879 GCTCAAGACGTCAGCTGTGCGCTCTGTGGTATGCTCTGTGTGTGTGTGTGTGTGTGTGTGT 938
Qy      1440 CCGGCTGAGCGCGCGCGGTGTCCTCGGTATCTCCCAACCGGAGCGCGCGCGCGCATCAAGAA 1499
Db      939 GCAAGTGAAGCTGCGCGCGAGTGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 990
Qy      1500 TGGGAGTGTGACAGGTGTCTTCAGTGTGCGAGTGTGAACAGGTATCTTACATGTGTGACA 1559
Db      991 -GATGAGTGAACCATTTGTATGAACAACGCGGTGAACAGGTATCTTACATGTGTGACA 1049
Qy      1560 CATGTGCTGTGCGCAACAGCTGCGCGCGCTGCAAGCGACGCGCGCGCGCGCTGTGCGCG 1619
Db      1050 CATGTGCTGTGCTAACCGCTGTGTGCGCTGCGCTCAAGAAAGCTGTGACGCGCTGTGCGCG 1109

```

QY 1620 CATCTCCGCGCGCCCATCAAGACGCTTAAGATCTACAG 1661
DB 1110 CATCTGCGCGCGCCCATCAAGACGCTTAAGATCTACAG 1151

RESULT 11

US-10-029-386-22715
Sequence 22715, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AEWICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 22715
LENGTH: 694
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL121929.17
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: NT HIT: g114746644, EVALU0 0.00e+00
OTHER INFORMATION: SWISSPROT HIT: P29503, EVALU0 3.00e-10
OTHER INFORMATION: EST_HUMAN HIT: AW340823.1, EVALU0 0.00e+00
US-10-029-386-22715

Query Match 11.1%; Score 186.6; DB 13; Length 694;
Best Local Similarity 58.1%; Pred. No. 2.8e-32;
Matches 349; Conservative 0; Mismatches 249; Indels 3; Gaps 1;

QY 681 CAGAGCTGAGAAACAACAGAGTGTGTGCAAGCTGGGCACTGGCGCTGGCGCGCC 740
DB 74 CAGAGGCGAGCAAGCGCGCTCTCGGTGAGCTTATGCACTCAAGTCCGCGGCGGCA 133
QY 741 GGGCCACCGCCAGCG 800
DB 134 CGGCGAGAGCG 193
QY 801 GGGCGCGCGGTGTGCG 860
DB 194 GCAAGCG 253
QY 861 GCCCG--ACGTGAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 917
DB 254 CGGCG 313
QY 918 CGGCG 977
DB 314 CGAGGCG 373
QY 978 GGTGGGCGCGTCCG 1037
DB 374 GGTCAAGCGCGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 433
QY 1038 CCCGCGCGGTCTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1097
DB 434 CCCG 493
QY 1098 AAGATCTGGGT 1157
DB 494 GGAATTTCTGGCGCGT 553
QY 1158 CAGCGTGGCT 1217
DB 554 GGTGTCAACG 613

QY 1218 GCTGTGCTGTGACACACAGCGAGCGCTCTGGCGCTTCTGCGCGTGGCGCGCGCTGC 1277
DB 614 GCTGTGCTGTGACAGCGCTCTGGCGAGCGCTTGTGATGTCTTTCGCGCTGTGACAGGACATCAC 673

QY 1278 G 1278
DB 674 G 674

RESULT 12

US-09-808-387-13
Sequence 13, Application US/09808387
Patent No. US2002012293A1
GENERAL INFORMATION:
APPLICANT: Kala Palm
APPLICANT: Tonis Timusk
TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
FILE REFERENCE: CEMRES.001A
CURRENT APPLICATION NUMBER: US/09/808,387
CURRENT FILING DATE: 2001-03-14
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
TYPE: DNA
ORGANISM: mouse
US-09-808-387-13

Query Match 10.9%; Score 182.4; DB 10; Length 876;
Best Local Similarity 58.8%; Pred. No. 2.3e-31;
Matches 315; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 99 GGTCTGTGGCGAGGCGCGCGCGCTTCCAGCGCGAGCGCAAGGCAAGACGTCGCGCTGA 158
DB 171 GCTCCGCGCAGCGCGCGCTTCTTCCACCCCACTAAGGCGCTCCGAGATCTCATGGA 230
QY 159 CGGCACTGTGCG 218
DB 231 CCTCAAGCAAGAGCGCGCTTCAAGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 290
QY 219 GCGGCGCATCGGCTGTACAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 278
DB 291 CGGCGCGGTCTCATCTACAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 350
QY 279 GAGCGCGCGCGCGCGCTTCCGCTTCAAGCGCGAGATCGTGTGTGTGTGTGTGTGT 338
DB 351 GAGCGCGCGCGCGCGCTTCCGCTTCAAGCGAGAGATCGTGTGTGTGTGTGTGTGT 410
QY 339 CATCCCAAGTACAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 398
DB 411 GCTGTCAAGTACAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 470
QY 399 GCCCGAAGACTGTGCG 458
DB 471 GCTGTGAGGAGTTTGCACAGAGGCAATCATCTCTTGTGTGTGTGTGTGTGTGTGT 530
QY 459 CCGTGTCTTACAGCGCGTACAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 518
DB 531 CCGTGTCTTACAGCGCGTACAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 590
QY 519 CGGCGCGCGCTGTGCGCTCATTTGATGTCTTACAGCGCGCGCGCGCGCGCGCGCGCG 578
DB 591 GGAAGCGCTGT 650
QY 579 GAGGCGCTTGT 634
DB 651 CAGCGAGCTGT 706

RESULT 13

```

US-09-808-387-19
; Sequence 19, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemline Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808,387
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 888
; TYPE: DNA
; ORGANISM: rat
US-09-808-387-19

```

Query Match 10.6%; Score 177; DB 10; Length 888;

Best Local Similarity 57.5%; Pred. No. 3.7e-30;

Matches 318; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

```

Qy 33 CCCGAGCCACCGCGCGCTCTGCGCAACCGCGCGCTGCGGCGCCCGCGCCCGAGCG 92
Db 105 CCCGATCCCTCTCAACGATGCAATCAAGAGAGAGATGCGCGCGCGCTGTCAG 164
Qy 93 ACGCCGCTCTGCGCGAGCGCGCGCGCTTCCACCGCGCAAGAGAGAGAGAGAGAG 152
Db 165 TGGGGGGCTCCCGCGCACACCGCTGCTTCCACCGCGCAAGAGAGAGAGAGAGAG 224
Qy 153 GCTGAGCGGCACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 212
Db 225 CATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284
Qy 213 CACGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 272
Db 285 CAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 344
Qy 273 TGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 332
Db 345 CTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 404
Qy 333 CCAGGACATCCCAAGTACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 392
Db 405 CGACTCACTGCCCAAGTACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 464
Qy 393 GGCACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 452
Db 465 AGCGTTCGCGAGGAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 524
Qy 453 CGCGCGCGCTGTTTCAACGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 512
Db 525 GGGCGGAGTCTTCAACGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584
Qy 513 CGTGGGCGCGCGCGCTCTGCGCGCTCATTTATGTCTACAGGACATCAAGAGAG 572
Db 585 AACGGCGGAGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCT 644
Qy 573 TCTGAGAGCGCGC 585
Db 645 GCTAGGAGAGAGCGC 657

```

```

RESULT 14
US-09-808-387-17
; Sequence 17, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemline Research

```

```

; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808,387
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: rat
US-09-808-387-17

```

Query Match 10.6%; Score 176.8; DB 10; Length 1035;

Best Local Similarity 57.4%; Pred. No. 4e-30;

Matches 319; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

```

Qy 33 CCCGAGCCACCGCGCGCTCTGCGCAACCGCGCGCTGCGGCGCCCGCGCCCGAGCG 92
Db 105 CCCGATCCCTCTCAACGATGCAATCAAGAGAGAGATGCGCGCGCGCTGTCAG 164
Qy 93 ACGCCGCTCTGCGCGAGCGCGCGCGCTTCCACCGCGCAAGAGAGAGAGAGAGAG 152
Db 165 TGGGGGGCTCCCGCGCACACCGCTGCTTCCACCGCGCAAGAGAGAGAGAGAGAG 224
Qy 153 GCTGAGCGGCACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 212
Db 225 CATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284
Qy 213 CACGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 272
Db 285 CAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 344
Qy 273 TGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 332
Db 345 CTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 404
Qy 333 CCAGGACATCCCAAGTACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 392
Db 405 CGACTCACTGCCCAAGTACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 464
Qy 393 GGCACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 452
Db 465 AGCGTTCGCGAGGAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 524
Qy 453 CGCGCGCGCTGTTTCAACGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 512
Db 525 GGGCGGAGTCTTCAACGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584
Qy 513 CGTGGGCGCGCGCGCTCTGCGCGCTCATTTATGTCTACAGGACATCAAGAGAG 572
Db 585 AACGGCGGAGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCT 644
Qy 573 TCTGAGAGCGCGCTTC 588
Db 645 GCTAGGCTTCACGATC 660

```

```

RESULT 15
US-09-808-387-11
; Sequence 11, Application US/09808387
; Patent No. US20020132293A1
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Tonis Timmusk
; APPLICANT: Cemline Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808,387
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 11
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: mouse
US-09-808-387-11

Query Match 10.5%; Score 176.4; DB 10; Length 1035;
Best Local Similarity 60.0%; Pred. No. 4.9e-30;
Matches 294; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

```
QY 99 GGTCTGGGCGAGGCGCGCGCTTCCACGGCGCAGGCCAAAGCAGAAAGCTGGCGCTGGA 158
   |||||
Db 171 GCTCCGGGCGACGCCGCTGCTTCCACCCCCACACTAAAGGCTCCAGATCCTCATGGA 230
   |||||
QY 159 CGGCGCATGCGCGCGCGCGGCGACACAGGGGCAACAGCTTGTGCAATGGGCTCACGTTCA 218
   |||||
Db 231 CTTAGCCACAAAGGCGGCTCAAGAGGCAAGGCCAGCTTGTGCAATGTCATCACTTCACTA 290
   |||||
QY 219 GCGGCCCATCCGCGCTGTACGAGCAGGTGCGGCTGCGCGCTGTGTGCGCGCGCTGCG 278
   |||||
Db 291 CCGCCCGGTGTCTATCTACGAGCAATCAGGCTGAAGATCACAGAGAGCAATGCTGTG 350
   |||||
QY 279 GAGCGGCGCGCTGCGCTTGCACCGCGCACGATCCGTGCTCATGAGCGCCAGGA 338
   |||||
Db 351 GAGCGGGGCGCTGCGCTTGCACCGAGCAAGGACCTTCCGCAATCCACCCGACTC 410
   |||||
QY 339 CATCCCAAGTACGCGCTGCGCGGACCTGTGCAAGCGCGCGGCTACTGGGCGCAAGGCACT 398
   |||||
Db 411 GCTGCCAAGTAGCGCTGCGCTGACCTGTGTCTCAAGTGTGCTTGTGGGCGCAAGCAATT 470
   |||||
QY 399 GCCCGAGAACTGCGCTGCGCGACAGGTGCTGCGCTACTGCGCGCGACCGCGCGCG 458
   |||||
Db 471 GCTGAGAGATTGCAACAGAGGCGCAACATATTGCTTGTGGTGAACAAGAGGCGG 530
   |||||
QY 459 CGTGTCTACAGCGTGAACGACGCGGAGCGGTCTTTCACCTGCGCGGTGCGCGTGG 518
   |||||
Db 531 CGTCTTCTACCGGATCAATGAGTCAAGCTGTATGCTTTCTTCAGTGGGCTCCGAGCGGT 590
   |||||
QY 519 CGGCGCGCTGCGCGCTCATTTGATGTCTACGGGATCACCGAGCAGAGTGAGCTTGA 578
   |||||
Db 591 GGAACCGCTGTGGGCGCTGTGAGCGTCTACGGGCTCACGCGGGGTGTCAAGCTGTAGG 650
   |||||
QY 579 GAGCGGCTTC 588
   |||||
Db 651 CTCACCATC 660
   |||||
```

Search completed: January 20, 2004, 20:10:06
Job time : 586 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 20, 2004, 15:07:52 ; Search time 3269 Seconds
(Without alignments)
12453.349 Million cell updates/sec

Title: US-09-808-387-21

Perfect score: 1675
Sequence: 1 atgggcaacacgctgcacg.....ctacagccatagcctagcc 1675

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_earth:*
2: em_earth:*
3: em_earth:*
4: em_earth:*
5: em_earth:*
6: em_earth:*
7: em_earth:*
8: em_earth:*
9: gb_earth:*
10: gb_earth:*
11: gb_earth:*
12: gb_earth:*
13: gb_earth:*
14: gb_earth:*
15: em_earth:*
16: em_earth:*
17: em_earth:*
18: em_earth:*
19: em_earth:*
20: em_earth:*
21: em_earth:*
22: em_earth:*
23: em_earth:*
24: em_earth:*
25: em_earth:*
26: em_earth:*
27: em_earth:*
28: gb_earth:*
29: gb_earth:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	517	30.9	573	B1848332
2	509.6	30.4	517	A1571301
3	492.8	29.4	644	AA850748
4	362.8	21.7	470	AM049442

Result No.	Score	Query Match Length	DB ID	Description
5	343.8	20.5	464	A113327
6	339	20.2	3212	AK078529
7	296.8	17.7	703	BU238931
8	288.6	17.2	377	A1579407
9	278.6	16.6	593	BB097995
10	275.4	16.4	392	AW529983
11	264.4	15.8	665	BM440106
12	263	15.7	304	BI404107
13	261.8	15.6	472	AA851139
14	242.2	14.5	311	A1577954
15	220.2	13.1	362	A2953956
16	210.6	12.6	355	A2341570
17	206	12.3	537	BI134290
18	199.2	11.9	527	BB649566
19	194.2	11.6	255	AA957126
20	194.2	11.6	318	AA489411
21	185.8	11.1	623	CA941769
22	185.8	11.1	626	BI712399
23	185.8	11.1	681	BI712399
24	185.8	11.1	715	U69206
25	182.6	10.9	578	BB656496
26	182.4	10.9	566	BU952430
27	182.4	10.9	588	BQ269642
28	181	10.8	570	CA948476
29	180	10.7	568	CA941491
30	178.4	10.7	984	BU452591
31	177.6	10.6	615	BI712632
32	175.8	10.5	589	CA771909
33	175.8	10.5	631	CA774789
34	175.4	10.5	2093	AK044464
35	172.2	10.3	661	BF147610
36	171.8	10.3	1060	AK010787
37	171.8	10.3	4034	AK046382
38	169.4	10.1	573	CA843851
39	169.4	10.1	645	BQ286290
40	167.4	10.0	431	BM724707
41	167.4	10.0	431	BM930417
42	167.4	10.0	447	BP591302
43	166.8	10.0	436	A1220894
44	163.2	9.7	559	CA941339
45	162.8	9.7	215	BB863961

ALIGNMENTS

RESULT 1
LOCUS B1848332
DEFINITION 470743 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION B1848332
VERSION B1848332.1 GI:15960851
KEYWORDS
SOURCE
ORGANISM Bos taurus (cow)
Bos taurus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
REFERENCE Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,B., Wray,J.E., White,J., Cho,J., Fahnensterg,G.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pettee,G., Holt,I., Karamycheva,S., Liang,P., Quackenbush,J. and Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
TITL B1848332
JOURNAL MEDLINE
PUBMED 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366

Fax: 402 762 4390
 Email: smilth@mail.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980504.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCACGACGACG
 Plate: 97 row: J column: 6
 Seq primer: ATTAGTACACTATAG.
 Location/Qualifiers

FEATURES

source

1..573
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_1ib="MARC 2B0V"
 /note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
 library made from pooled tissue from testis, thymus,
 semitendonsus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium."
 BASE COUNT 88 a 214 c 191 g 80 t

Query Match

Best Local Similarity 93.9%; Score 517; DB 12; Length 573;
 Matches 538; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

112 GCGCCGCGCTTCCACGCGGACGAGGCGCAAGAGCGTGGCTGAGCGGCACTGCGCC 171
 1 GCGCGCGCTTCCACGCGGACGAGGCGCAAGAGCGTGGCTGAGCGGCACTGCGCC 60
 112 GCGCGCGCTTCCACGCGGACGAGGCGCAAGAGCGTGGCTGAGCGGCACTGCGCC 231
 61 CGAGCCACGCGGACGAGGCGCTTCCACGCGGAGCGTGGCTGAGCGGCACTGCGCC 120
 232 CTGTACGAGCAGGTGGCTGGCGCTGTGGCGCTGTGGCGCTGTGGCGCTGTGGCGCTG 291
 121 CTGTACGAGCAGGTGGCTGGCGCTGTGGCGCTGTGGCGCTGTGGCGCTGTGGCGCTG 180
 292 CGCTTGGCTTCCACGCGGACGAGTCCGTGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 351
 181 CGCTTGGCTTCCACGCGGACGAGTCCGTGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 240
 352 GCGTGGCGGAGCTGTGTACGCGGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 411
 241 GCGTGGCGGAGCTGTGTACGCGGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 300
 412 GCGTGGCGGAGCTGTGTACGCGGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 471
 301 GCGTGGCGGAGCTGTGTACGCGGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 360
 472 GTGAAGCAGCGGAGCGGCTGTGTACGCGGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 531
 361 GTGAAGCAGCGGAGCGGCTGTGTACGCGGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 420
 532 GCGTGGCGGAGCTGTGTACGCGGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 591
 421 GCGTGGCGGAGCTGTGTACGCGGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 480
 592 GACAGCGTGAAGCGCGGCGGCTGTATGAGCGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 651
 481 GACAGCGTGAAGCGCGGCGGCTGTATGAGCGCGGCTGTATGAGCGCGGCAAGATCCCAAGTAC 540
 652 AAGCAGCAGCGGCGGCAAGTTCGACCAAGAG 684
 541 AAGCAGCAGCGGCGGCAAGTTCGACCAAGAG 573

RESULT 2
 A1571301/c
 LOCUS A1571301

517 bp

mRNA

linear

EST 12-MAY-1999

DEFINITION tna4e06.x1 NCI CGAP Brn25 Homo sapiens cDNA clone IMAGE:2170498 3'
 similar to TR:Q24746 Q24746 NEURALIZED PROTEIN-1, mRNA sequence.
 ACCESSION A1571301
 VERSION A1571301.1 GI:4534675
 KEYWORDS EST.

SOURCE

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 517)
 NCI/NIHDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute / National Institute of Neurological
 Disorders and Stroke, Brain Tumor Genome Anatomy Project
 (CGAP/BrTAP), Tumor Gene Index

JOURNAL

Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
 Ph.D.

COMMENT

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMW at:
 www.bio.liml.gov/bbrp/image/image.html
 Insert Length: 956 Std Error: 0.00
 Seq primer: -40UP from Glibco
 High quality sequence stop: 435
 POLYA=No.

FEATURES

source

1..517
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_1ib="IMAGE:2170498"
 /tissue_type="neoplastic oligodendroglioma"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pRT73D-Pac (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Bco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer (5'
 TGTACCACTGTGAAGTGAAGCGGCGGCAATGTTTCTTTTCTTTTCTTTTCTTTTCTTTT
 T 3'); double-stranded cDNA was ligated to Bco RI
 adaptor (Pharmacia), digested with Not I and cloned into
 the Not I and Bco RI sites of the modified pRT73 vector.
 Library is normalized, and was constructed by Bento
 Soares and M. Fatima Bonaldo."

BASE COUNT 76 a 172 c 190 g 78 t 1 others

Query Match 30.4%; Score 509.6; DB 9; Length 517;
 Best Local Similarity 99.0%; Pred. No. 1.9e-78;
 Matches 512; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

214 ACGCAGCGGCGGCTGTATGAGCAGGTGGCGCTGTGGCGGCTGTGGCGGCTGTGGCGGCT 273
 517 ACGCAGCGGCGGCTGTATGAGCAGGTGGCGGCTGTGGCGGCTGTGGCGGCTGTGGCGGCT 458
 274 GCGTGGAGCGGCGGCTGTGGCGGCTGTATGAGCAGGTGGCGGCTGTGGCGGCTGTGGCGGCT 333
 457 GCGTGGAGCGGCGGCTGTGGCGGCTGTATGAGCAGGTGGCGGCTGTGGCGGCTGTGGCGGCT 398
 334 CAGGACATCCCAAGTACGCGGCGGCTGTATGAGCAGGTGGCGGCTGTATGAGCAGGTGGCGGCT 393
 397 CAGGACATCCCAAGTACGCGGCGGCTGTATGAGCAGGTGGCGGCTGTATGAGCAGGTGGCGGCT 338
 394 GCACTGCCGGAACCTGGCGGCTGTATGAGCAGGTGGCGGCTGTATGAGCAGGTGGCGGCT 453
 337 GCACTGCCGGAACCTGGCGGCTGTATGAGCAGGTGGCGGCTGTATGAGCAGGTGGCGGCT 278
 454 GCGCGGCTTTCACAGGTGAAGCAGCGGCGGCGGCTGTATGAGCAGGTGGCGGCTGTATGAGCAGGTGGCGGCT 513

Db	277	GGCCGCGGTTCTACAGAGTGAACGACGCGGACCGGTCCTTCCACTGCGGCGTGACC	218
Qy	514	GTGGGCGGCGCCGCTCTGTGGGCGCTCAATTATGTCTTACGGCATATCCAGACGAGGTGCAGTT	573
Db	217	GTGGGCGGCGCCGCTCTGTGGGCGCTCAATTATGTCTTACGGCATATCCAGACGAGGTGCAGTT	158
Qy	574	CTGAGAGACGCGCTTCTGCTGACAGCGTGAACGCGCGCGCGCGCTCAGCAGGCGCGCTTACG	633
Db	157	CTGAGAGACGCGCTTCTGCTGACAGCGTGAACGCGCGCGCGCGCTCAGCAGGCGCGCTTACG	98
Qy	634	GCTTGCTCTCCGCGCCAGCAGCCAGACCGCGGCCAACTTGCACAAACGAGCTCGAGAC	693
Db	97	GCTTGCTCTCCGCGCCAGCAGCCAGACCGCGGCCAACTTGCACAAACGAGCTCGAGAC	38
Qy	694	AACCAAGTGATGGCCCAAGCTGCGGCCCACTGCGCGCTGG	730
Db	37	AACCAAGTGATGGCCCAAGCTGCGGCCCACTGCGCGCTGG	1
RESULT 3			
LOCUS	AA850748	644 bp	mRNA linear EST 30-APR-1998
DEFINITION	EST193516 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone		
ACCESSION	AA850748		
VERSION	AA850748.1	GI:2938288	
KEYWORDS	EST.		
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 644) Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.		
TITLE	Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat Gene Index		
JOURNAL	Unpublished		
COMMENT	Contact: Lee, NH The Institute for Genomic Research 9712, Medical Center Drive, Rockville, MD 20850, USA Tel: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@igr.org Seq primer: M13-21.		
FEATURES	Location/Qualifiers		
SOURCE	1..644 /organism="Rattus sp." /mol_type="mRNA" /db_xref="ATCC (Inhost):2009149" /db_xref="ATCC (Inhost):2009149" /db_xref="taxon:10118" /clone.jlb="ROVAK34" /clone.jlb="Normalized rat ovary, Bento Soares" /note="Organ: ovary; Vector: pT7T3Pac; Site_1: EcoRI; Site_2: NotI"		
BASE COUNT	87 a 218 c 226 g 113 t		
ORIGIN			
Query Match	29.4%	Score 492.8	DB 9; Length 644;
Best Local Similarity	85.6%	Pred. No. 1.5e-75;	
Matches 548; Conservative	0;	Mismatches 92;	Indels 0; Gaps
Db	953	GAGCCGAGAGAGCTCTTCTGCTGAGAGTGGGCGCGCTCCGAGGCTGCGCGCGCGCGCGCGC	1012
Qy	5	GCCCGGGGAGAGCTCTGCTGCTGAGAGTGGGCGCGCGCTCCGAGGCTGCGCGCGCGCGCGC	64
Db	1013	TGGCTTGGCATCATCGTGTGTCGACCCGGGCGTGTACAGGCTTACAGGCTGCGCGCGC	1077
Qy	65	TGGCTTGGCATCATCGTGTGTCGACCCGGGCGTGTACAGGCTTACAGGCTGCGCGCGC	124
Db	1073	ACCAAGAGCGGTGTCTTCCACCGCAAGAGTACTGGGTGGTGGCGCGCGCGCGCGCGCGTGC	1133
Qy	125	ATCCCGCTGCGTGTCTGAGCCCGCAAGATATCTGGGTGGTGGCGCGCGCGCGCGCGCGTGC	184

[illegible]

Db		181	GGGCCCCGTGCCCAAGCGAAGGCGACACTCAGCTTTCAGCTGCACCCGGGCGGGAAGCTC	240
Oy		1183	CTCCTGGGACATCAAAGGCGCTCCGCGCGCGCCCTGTGTGCTGTGACACCAACGACGAGCG	1242
Db		241	CTGCTGGGGGATGAACGGGCGCGCCGCGGGGAGACGTTTGCTGTGCGTAGAACACTCCGACGGCG	300
Oy		1243	CTCTGGGCGCTTCTTTCGCGCTGCGCGCGCGCGCGCTCCGCGGCGACGTCGCTCTCTGTCGTAACC	1302
Db		301	CTCTGGGCGCTTCTTTCGCTGTGCGCGGTGTGTGTGGCGGGGTACGTCGCTCTCTCTGGGCAACC	360
Oy		1303	CTGACATCCAGCACCCCTGCGACACGACGACTCATCAGGGTCCCTCAGCGGCTCCCAAGACAT	1362
Db		361	GTCACATCCGATCTCTGAGGCGCACACTCATCAGGGTCTTCAGTGGCTCTCAGAGATGAC	420
Oy		1363	AGTAAATTCAGATTATGACCTTCAGTGTCAACC	1393
Db		421	AGCGATTCGACATGACCTTCGGGGGTCAACC	451
RESULT 6				
AKO78529				
LOCUS		3212 bp	mRNA	linear
DEFINITION			Mus musculus 12 days embryo female mullerian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:6820434L06 product:hypothetical protein, full insert sequence.	
ACCESSION		AKO78529		
VERSION		AKO78529.1	GI:26097916	
KEYWORDS		HTC; CAP trapper.		
SOURCE		Mus musculus (house mouse)		
ORGANISM		Mus musculus		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS		Carninci,P. and Hayashizaki,Y.		
TITLE		High-efficiency full-length cDNA cloning		
JOURNAL MEDLINE		Meth. Enzymol. 303, 19-44 (1999)		
PUBMED		99279253		
REFERENCE		10349636		
AUTHORS		2		
TITLE		Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
JOURNAL MEDLINE		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
PUBMED		Genome Res. 10 (10), 1617-1630 (2000)		
REFERENCE		20499374		
AUTHORS		11042159		
TITLE		Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komono,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Obara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuzura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kiya,A. and Hayashizaki,Y.		
JOURNAL MEDLINE		Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipicillary sequencer		
PUBMED		Genome Res. 10 (11), 1757-1771 (2000)		
REFERENCE		20530913		
AUTHORS		11076861		
TITLE		4		
JOURNAL MEDLINE		Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Aikawa,T., Hara,A., Fubunishi,Y., Komono,H., Adachi,J., Fukuda,S., Aizawa,K., Irawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanka,I., Saito,T., Okazaki,Y., Gyojohri,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavola,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochwa,H., Kuell,P., Lewis,S., Matsuo,Y., Nikaido,T., Pesole,G., Quackenbush,J., Schriml,L.M., Stuhli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sekai,K., Okido,T., Furuno,M., Arano,H., Baldarelli,P., Barsh,G., Blake,J., Boftelli,D., Bojunga,N., Carinini,P., de Bonaldo,M.F., Brownstein,M.J., Bull,C.,		

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 PUBLISHED
 REFERENCE
 AUTHORS
 5
 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 3212)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furumai, M., Hanagata, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-research.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 COMMENT
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.
 FEATURES
 SOURCE
 1..3212
 location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM.DB:6820434L06"
 /db_xref="taxon:10090"
 /clone="6820434L06"
 /sex="female"
 /tissue_type="mullerian duct includes surrounding region"
 /clone_id="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="12 days embryo"
 1..3212
 /note="hypothetical protein (evidence: decoder, longest-ORF)"
 BASIS COUNT
 562 a 993 c 943 g 714 t
 ORIGIN
 Query Match 20.2% Score 339; DB 11; Length 3212;
 Beat Local Similarity 56.3% Pred. No. 5.6e-45;
 Matches 794; Conservative 0; Mismatch 555; Indels 62; Gaps 6;
 275 GCTGACCGGCGCGCTGCGCTTCACCGGACGATCGTGCCTCATGAGCGCC 334
 51 GCTGACCGGCGCGCTTCGACTTGGCTTCACACAGACCTTCCCGCATCCACCCG 110

QY 335 AGGACATCCCAAGTACGCTGCGCCGAGACCTGTGACAGCGGCGGAGCTACTGAGCCCAAG 394
 DB 111 ACTCGGTGCGCAAGTACGCTGCGCTGACCTGTGTCTCAAGATGAGCTTCTGAGCCAAAG 170
 QY 395 CACTGCGGAGAACTGGGCGTGGCGGACAGCGGTGCTGCTACTGAGGCGGACCGGACG 454
 DB 171 CATTTGCTGAGAGTTTGGCCAAAGAGGCAACATCATTTGCTTCTGGGTGACAAAGAG 230
 QY 455 GCGCGGTGTTCTACAGCTGGAACAGAGGAGCGGAGCTTCTCACTGCGGCGGTGGCGG 514
 DB 231 GCGGCTCTTCTACCGGATCAATGAGTCAGCTGTATGCTTCTTCTTCAAGTGGGATCCGA 290
 QY 515 TGGGCGGCGGCTCTGGGCGCTGATGATGTCTACGCAATCAACGAGAGTGCAGCTTC 574
 DB 291 CGGTGAGCCCGCTCTGGGCGCTGTGAGACCTCAACGCGCTCAAGCGGGGTGTCACTGG 350
 QY 575 TGGAGAGCGGCTTCTGCTGACAGCTGACAGCGCGCGGCTCAGCCAGGCGGCTTCAAGG 634
 DB 351 TAGACAGC-----GAGCTGTGTGCTGCCAGCTGCTGCGGCGGCTCTCTTCAAGC 401
 QY 635 CCGCTGCGCGCGCCAGAGCCAGAGCCGCGGCACTTCAACAAACAGCTGAGAGAA 694
 DB 402 CGCTGCGGCGGCGCTGCTGCGGTGCGAGCGGAT-----GAGCGCG 443
 QY 695 ACCAGGTGTGCGCAAGCTGGGCGCACCTGCGCTGGGCGCGCGCGCGGCCCAACGCGAG 754
 DB 444 GCGTGTGAGTGTGCGCTGTGACACTCAAGTGCAGGAGCG--ACGGAGAGAGCGGCGGT 501
 QY 755 CCGACCGCGCGCGCGCGCGCTTCTGCGGCGGCGCGGCTGAGCGCGCGCGCGCGCTCT 814
 DB 502 CACCGGCTGCGCGCTGCGCGCTGCGCGAGAACTGCTCAATTTCAAGCACAGCGCGCGCG 561
 QY 815 CCGCGGCGCTACTGAGAGCGCGAGCTGCGCTT-----CGACGAAACAGCGCGCGCGAGC 868
 DB 562 TGCCTGCGGCGCTGAGAGCGGAGCTGCGCTTCTTCAAGCGCGCTTCTGCGCGCGCGAGC 621
 QY 869 TGAAGCTGTGCGCGCGAGCGGAGAGTGTGCGGACCGCGCGCGCGCGCGCGCGCGAGC 928
 DB 622 TCGGATGTGTGAGAGAGAGAGAGAGCGGTGCGCGCTGAGAGAGCGGCGCGAGAGCGCGCG 681
 QY 929 TGTCTTCTTCTGAGAGCGCGCGCTGCGCGCGCGCGAGAGCTTCTTCTGAGAGTGGCGCT 988
 DB 682 TCTCTTCTTCAACACCGCGCTGTGTGCGCGCGAGACCATCTTCAATCAAGTCAAGCGCT 741
 QY 989 CCGGCGGTGCGCGCGCGCGCGCGCTTCTGCGGATCAAGTGTGAGAGCGCGCGCGCTG 1048
 DB 742 CCGGCGGCGGCGAGAGCGGCGCGCGCTGCTTCTTCTGCGGATCAACAGTGTGAGAGCTG 801
 QY 1049 TACGCGCCAAAGAGCTGCGCGCGCGAGCCAGAGCGGCTGCTCAACCGCAAGAGTACTGG 1108
 DB 802 TGGGCGCGCGCGAGAGCTGCTTCTTCAAGCGCGAGCGCTGTGAGAGCGGAGAGTTCGG 861
 QY 1109 TGTGTGCGCGCGCGCGCGCGCGCTGCGAGCGCGCGCGAGCGGCTTCAAGCTTCAAGCG 1168
 DB 862 CCGTGTGTGTGCGCGCGCGCGCTTGTGACAGCGCGAGCATCTGAGGCTGTGTGTCAAGC 921
 QY 1169 CCGGCGGCGAGAGTGTCTTGTGAGATCAAGAGGCGTCCGCGCGCGCGCTGCTGTGAGCTG 1228
 DB 922 CGGAGAGAGAGCTGAGCTTCAAGTCAACAGCGCGCGCGCGCGAGTGTGAGCTGTGCGTGG 981
 QY 1229 ACAACAGAGAGGCGCTTGTGAGCTTCTTCTGCGCGGTGCGCGCGGCGTGTGCGAGCTGC 1288
 DB 982 ATGCGCTGAGAGCGCGCTGTGATGTCTTCAAGCGCTGAGCGGAGCATCAAGAGTTCGCG 1041
 QY 1289 GTTCTCTGAGTACCTGTGAGTCAAGCGCTGCGAGAGAGTCAATCAAGGCT----- 1339
 DB 1042 TCTTCTGAGTCAACATCATATAGTGAAGCGGAGTGGCGCATCTCTCTCTGCTCAAGCTG 1101
 QY 1340 -----CCCTCAGCGGCTCCAGAGAGATAGTATGATGATGATGATGATGATGATGATG 1390
 DB 1102 CCACTCCAACTCAACAGAGTGGCGGTGTGAGCATCCGCTCTTGTGAGAGCGCGCTGTGAGC 1161
 QY 1391 ACCAGTCTCTCTGCGGAGTGTGAGTCACTCTGTGTGAGAGCGCGCGCGAGCTTCCCGCTG 1450

DB 1162 GCGGTTCGTGAGCGCGCTAGAGTGTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1221
 QY 1451 CCGCGGTGTCCCGGTGTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1510
 DB 1222 TGCAGAGTCAACAG 1272
 QY 1511 CCGGTGTCTTCAAG 1570
 DB 1273 CCATTGTCTAAG 1332
 QY 1571 GCGAGAGTGTGAG 1630
 DB 1333 GCTACTCTGTGTGAG 1392
 QY 1631 GCGCAGTCAAG 1661
 DB 1393 GCGCAGTCAAG 1423

RESULT 7
 BU238931 703 bp mRNA linear EST 26-NOV-2002
 LOCUS 603322904P1 CSBQCHN33 Gallus gallus cDNA clone ChsBT251g2 5', mRNA
 DEFINITION
 ACCESION BU238931
 VERSION BU238931.1 GI:25484991
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 703)
 Boardman, P. E., Sanz-Bzquez, J., Overton, I. M., Burt, D. W., Bosch, E.,
 Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 12445392
 MEDLINE
 PUBMED
 COMMENT Contact: Simon Hubbard
 Department of Biomedical Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..703
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Compton Line 151"
 /db_xref="taxon:9031"
 /clone="ChsBT251g2"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_1lb="CSBQCHN33"
 /note="Organ: liver; Vector: pBluescript II KS(+); Site_1:
 EcoRI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunt-ended, ligated to NotI adapters, digested with EcoRI
 , size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

BASE COUNT 150 a 228 c 202 g 122 t 1 others
ORIGIN

Query Match 17.7%; Score 296.8; DB 13; Length 703;
Best Local Similarity 72.8%; Pred. No. 9.9e-42;
Matches 456; Conservative 0; Mismatches 152; Indels 18; Gaps 5;

QY 1 ATGGGCAACACGGTGCACCGGACCTTGCAGACCCGAGCCACCGGCGCGCTCTCTGACC 60
DB 77 ATGGGCAACACGGTGCACCGGACCTTGCAGACCCGAGCCACCGGCGCGCTCTCTGACC 136
QY 61 ACCGGCGCGTGTGGGGCC-----CGGCCCCGAGCCAGCCCGGCTCTGAGC--- 108
DB 137 AGCCCTCCCTACTAGACGCTGCCCAACGACCCAGCCAGCCGCTCAGTGTCTGCTGTC 196
QY 109 ---GAGGCGCGCGCTTCCAGCGGAGGCAAGAGGCAAGAGCGTGGCTGAGACGGCCAC 165
DB 197 ACGGAGCGCGTGTGGCTTCCAGCTCCAGCCGCAAGGCAAGAGCGTGGCTGAGACCCAC 256
QY 166 TCGCGCGCGGCGCAACGCGGCAACAGCTTCTGCAATGGCGTCAAGCTTCAACGAGCGGCC 225
DB 257 TCCGAGCGGCGCAACGAGGAGGAGCAAGCTTCTGTAACGGATCACCTTCAACGAGCGGCC 316
QY 226 ATCCGCGTGTACGAGAGGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285
DB 317 ATCCAGCTGTACGAGAGGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376
QY 286 GCGCGCGCTTCTGCGCTTCAACCGGCGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTG 345
DB 377 GCGCGCGCTTCTGCGCTTCAACCGGCGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTG 436
QY 346 AAGTACGCGCTGCGCGGACCTGTGTCAACGCGCGCGGCTACTGCGGCGCAAGGCACTGCG 405
DB 437 AAGTACGCGCTGCGCGGACCTGTGTCAACGCGCGCGGCTACTGCGGCGCAAGGCACTGCG 496
QY 406 AACTGCGCGCTGCGCGGACCTGTGTCAACGCGCGCGGCTACTGCGGCGCAAGGCACTGCG 465
DB 497 AGTGTGCTGTACGAGGAGCAACGCTGCGCTGCTGCGCTGCTGCGCTGCTGCGCTGCTG 555
QY 466 TACACGCGTGAACGAGGCGGAGCGGCTGCTTCACTGCGCGCGGCTGCGGCGCGCGCG 525
DB 556 TACACGCGTGAACGAGGCGGAGCGGCTGCTTCACTGCGCGCGGCTGCGGCGCGCGCG 614
QY 526 CTCTGGGCGCTCATGTATGTCTACGCAATCAACCGACG-AGGTGACAGCTTCTGAGAGCGC 584
DB 615 CTCTGGGCGCTCATGTATGTCTATGGAATCAACCGAGGAGGCAATCTAGACGAGAT 674
QY 585 CTTCGCTGACACGCTGACGCGCGCGC 610
DB 675 GTTTGACAGACCATGACCATGCGCC 700

RESULT 8
A1579407 377 bp mRNA linear EST 05-APR-1999

LOCUS
DEFINITION UI-R-AG0-wy-b-01-0-UI.61 UI-R-AG0 Rattus norvegicus cDNA clone

ACCESSION A1579407
VERSION A1579407.1 GI:4563783
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 377)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MBRRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Clones will be available through
Research Genetics (www.reagen.com) The following repetitive
elements were found in this cDNA sequence: 39-65,
>GC-rich#low complexity
Seq primer: M13 Forward.
Location/Qualifiers
1..377
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-AG0-wy-b-01-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-AG0"
/note="Vector: pUT19D-Pac (Pharmacia) with a modified
polylinker Site_1: Not 1; Site_2: Eco RI; The UI-R-AG0
library is a non-normalized library constructed from 13
dpc rat ventricle. The tag is a string of 6 nucleotides
present between the Not I site and the oligo-dT track.
Lennon and Soares, Genome Research 6: 791-806, 1996.
Tissue provided by Jim Lin, Department of Biology,
University of Iowa."

BASE COUNT 36 a 130 c 149 g 61 t 1 others
ORIGIN

Query Match 17.2%; Score 288.6; DB 9; Length 377;
Best Local Similarity 85.4%; Pred. No. 2.5e-40;
Matches 321; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 943 GCGCGCTGCGCGCGCGGAGAGCTTCTGCTGAGAGTGGGCGCGCGCGCGCGCGCGCGCG 1002
DB 1 CGGCGCTGCGCGCGCGGAGAGAGCTTGTGCTGAGAGTGGGCGCGCGCGCGCGCGCGCG 60
QY 1003 CCGCGCGCTGCGCGCTTCTGCGATCATGCTGTGCAACCGCGGAGTCTACGCGCCACGAG 1062
DB 61 CCGCGAGCTGTGCGCTTCTGCGATCATGCTGTGCAACCGCGGAGTCTACGCGCCATCGAG 120
QY 1063 CTGCGCGCGCGAGCGCGCGCTGTGTCAACCGCGGAGTCTACGCGCCATCGAG 1122
DB 121 CTGCGCGCGCGAGCTGCGCTGTGTCAACCGCGGAGTCTACGCGCCATCGAG 180
QY 1123 GGGCGCGCGAGCGCGCGGAGCGCGCTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 1182
DB 181 GGGCGCGCGAGCGCGCGGAGCGCGCTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 240
QY 1183 CTCTGGGCGATCAACGCGCGCTGCGCGCGCGCGCTGCTGTGTGAGCAACGAGGCG 1242
DB 241 CTGCGTGGGAGTGAACGCGCGCGCGCGCGGAGCGCTTGTGTGTGAGCAACCTCGAGGCG 300
QY 1243 CTCTGGGCGCTTCTTCTGCGCGTGTGTGCGCGCGCGCTGCGCGCGCGCGCGCGCGCT 1302
DB 301 CTCTGGGCGCTTCTTCTGCGCGTGTGTGCGCGCGCGCTGCGCGCGCGCGCGCGCGCT 360
QY 1303 CTGAGTCCAGCGCTG 1318
DB 361 GTACAGTCCGCTCTG 376

RESULT 9
B8097995 593 bp mRNA linear EST 12-JUN-2000
LOCUS
DEFINITION UI-R-C4-anc-b-05-0-UI.61 UI-R-C4 Rattus norvegicus cDNA clone
UI-R-C4-anc-b-05-0-UI 3', mRNA sequence.

ACCESSION BE097995.1 GI:8488889
 VERSION
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 593)
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Oligo-dt track not found. Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 39-65,
 >GC rich#Low complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source 1..593
 /location/Qualifiers
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C4-anc-b-05-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-C4"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C4
 library is a subtracted library of a series, ultimately
 derived from a mixture of tissues from rat placenta,
 adult lung, brain, liver, kidney, heart, spleen, ovary,
 muscle, and 8, 12 and 18-day embryos. For a detailed
 description of the library from which this clone was
 derived, please visit our web site at
 ratest.eng.uiowa.edu. This procedure has been previously
 described (Ronaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_SEQ=None found"
 BASE COUNT 94 a 162 c 200 g 136 t 1 others
 ORIGIN
 Query Match 16.6%; Score 278.6; DB 10; Length 593;
 Best Local Similarity 86.3%; Pred. No. 1.3e-38;
 Matches 308; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 943 CGCCCGCTGCGCCCGCGAGAGCTTCTGTGTAGAGTGGCCGCTCCGGGCTGGCGCGC 1002
 Db 1 CGCCCGCTGCGCCCGCGAGAGCTTGTGTGTAGAGTGGCGCGCCCGGGCTGGCGCGC 60
 QY 1003 CCGCGCGCTGCGCTTCTGCGATCACTCTGCGATCTGCGCGCTGCGCGATCCGAG 1062
 Db 61 CCGCGAGCTGTGCGCTTCTGCGATCACTCTGCGATCTGCGCGCTGCGCGATCCGAG 120
 QY 1063 CTGCGCGCGAGCCGAGCGCGCTGTGCGACCGCAAGAGTACTGGGTGTGGCGCGCC 1122
 Db 121 CTGCGCGCGATCCCGCTGCGCTGTGCGACCGCAAGAGTACTGGGTGTGGCGCGCC 180
 QY 1123 GGGCGCGTGTGCGAGCGCGCGAGCGCGCTGAGCTTCAGCTGTGGCGCGCGCGAGCTG 1182

Db 181 GGGCGCGTGTGCGAGCGAGCGAGCGACTTACGCTGTGCGACCGCGCGCGAGCTTC 240
 QY 1183 CTCTGGGATCAACGGGCGTTCGGCGCGCGCTGTGTGTGTGCGACACGAGCGG 1242
 Db 241 CTGCTGGCGGTGAACGGGCGCGCGCGCGAGACCTTGTGTGTGTGAGACCTTCGAGCG 300
 QY 1243 CTGTGGGCTTCTTGTGCGCGGTGTGCGCGCGCGGTGTGCGCGCGAGCTGCTCTGGGT 1299
 Db 301 CTGTGGGCTTCTTGTGCGCGGTGTGCGCGCGGTGTGCGCGGTGAGCTGCTCTGGGT 357

RESULT 10
 AM529983
 LOCUS 392 bp mRNA linear B87 06-MAR-2000
 DEFINITION UI-R-C4-alc-f-06-0-UI.81 UI-R-C4 Rattus norvegicus cDNA clone
 UI-R-C4-alc-f-06-0-UI 3', mRNA sequence.
 ACCESSION AM529983
 VERSION AM529983.1 GI:7172397
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 392)
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Oligo-dt track not found. Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 39-65,
 >GC rich#Low complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source 1..392
 /location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C4-alc-f-06-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-C4"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C4
 library is a subtracted library of a series, ultimately
 derived from a mixture of tissues from rat placenta,
 adult lung, brain, liver, kidney, heart, spleen, ovary,
 muscle, and 8, 12 and 18-day embryos. For a detailed
 description of the library from which this clone was
 derived, please visit our web site at
 ratest.eng.uiowa.edu. This procedure has been previously
 described (Ronaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_LIB=UI-R-C4
 TAG_TISSUE=corpus-strictum
 TAG_SEQ=CTAGG"
 BASE COUNT 38 a 132 c 156 g 66 t

Query Match 16.4%; Score 275.4; DB 9; Length 392;
 Best Local Similarity 85.7%; Pred. No. 4.7e-38;
 Matches 306; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 943 CGCCCGTCGCGCCCGCGAGAGCTCTTCGTGAGGAGGCGCGTCCGAGGCTGAGCGC 1002
 Db 1 CGGCGCTGCGCGCCCGGAGAGGCTGCGTGAAGTGGGCGCCCGGCGCTGGCGCG 60

Qy 1003 CCGGCGCGCTGCGCTTCGTCATCACTGTCGCAAGCCGCGGCTGTCACGCGCAACGAG 1062
 Db 61 CCGCGAGCTGTCGCTTCGTCATCACTGTCGCAAGCTGTCGCGCTGCGCGCAATTCGAG 120

Qy 1063 CTGCGCGCGCAAGCCGAGAGCTGTCGCAAGGAGTACTGAGTGTGTGGCGCGCC 1122
 Db 121 CTGCGCGCGCAAGCTGTCGTCGTCGTCGCAAGGAGTACTGAGTGTGTGGCGCGCC 180

Qy 1123 GGGCGCGTCGCGAGCGCGCGCGAGCTGTCATCACTGCGCGCGCGCGAGAGCTG 1182
 Db 181 GGGCGCGTCGCGAGCGCGCGAGAGCTGTCATCACTGCGCGCGCGAGAGCTG 240

Qy 1183 CTCTGGGATCAAGCGGCGCTGCGCGCGCGCGCTGTCGTCGTCGACACGAGGCG 1242
 Db 241 CTCTGGGATCAAGCGGCGCTGCGCGCGCGCGCTGTCGTCGTCGACACCTCGAGGCG 300

Qy 1243 CTCTGGGCTCTCTTCGCGCGCGCGCGCGCGCGCTGTCGTCGTCGTCGTCGTCGTC 1299
 Db 301 CTCTGGGCTCTCTTCGCGCGCGCGCGCGCGCGCTGTCGTCGTCGTCGTCGTCGTC 357

RESULT 11

BM440106

LOCUS

BM440106 665 bp mRNA linear EST 01-FEB-2002
 pgrin.pk002.c6 Normalized Chicken Reproductive Tract cDNA library
 (pgrin) Gallus gallus cDNA clone pgrin.pk002.c6 5' similar to
 gb|AAK84420.1|AF400063_1 (AF400063) neutralized 1 [Mus musculus],
 mRNA sequence.

ACCESSION

BM440106

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
 University of Delaware and INRA, Tours-Poultry Unit Project
 Contact: Larry A. Cogburn
 University of Delaware
 Townsend Hall, Newark, DE 19717, USA
 Tel: 302-831-1335
 Fax: 302-831-2832
 Email: cogburnudel.edu, www.chickent.udel.edu.

FEATURES

SOURCE

1. 665
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Commercial broiler and layer"
 /db_xref="taxon:9031"
 /clone="pgrin.pk002.c6"
 /sex="Male and Female"
 /tissue_type="Testis, ovary and oviduct"
 /dev_stage="Various stages; embryonic, post-hatch, immature
 and sexually-mature"
 /lab_host="E. coli EMDH10B"
 /clone_lib="Normalized Chicken Reproductive Tract cDNA
 Library (pgrin)"
 /note="Vector: PCMVSPORT6; Library made from three total
 RNA pools from each tissue (testis 25%, ovary 25%, and
 oviduct 50% of final RNA pool); Single pass sequencing
 from 5'-end"

BASE COUNT 146 a 210 c 174 g 135 t
 ORIGIN

Query Match 15.8%; Score 264.4; DB 12; Length 665;
 Best Local Similarity 63.4%; Pred. No. 3.8e-36;
 Matches 425; Conservative 0; Mismatches 226; Indels 9; Gaps 1;

Qy 557 CCGACGAGTGTGAGCTTCTGAGAGCGCTTCGTGACAGCTGACCGCGCGCTTCA 616
 Db 5 CCGACGAGAGCTTCTGAGAGCGCTTCGTGACAGCTGACCGCGCGCTTCA 64

Qy 617 GCCAGGCGCTTCAGGCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 676
 Db 65 GCACTGCGCGCTTCAGGCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 124

Qy 677 ACAACGAGCTGAGAGCAACCAAGTGTGTGCAAGCTGTGAGCACTGTGAGCGCGCG 736
 Db 125 ACAATGAACTGTGAGAGCAACCAAGTGTGTGCAAGCTGTGAGCACTGTGAGCGCGCG 184

Qy 737 CCGCGGCG 796
 Db 185 TGCCAGGACTGTGTGACAGACA-----CCACATCATCTCTGTGCGCCCAACCGCGC 235

Qy 797 GCCCG 856
 Db 236 GGCACGCTGTTCAGAGGAGTCCCGGCTTCTGTGACACAGACTGTGACATTCACCTTACC 295

Qy 857 GCGGCG 916
 Db 296 ATGACCGCGACATCATCTTCTTCTGAGAGCGAGTGTGTGAGCGAGCACTGTGAGAGGA 355

Qy 917 GCGGCG 976
 Db 356 GCATATGAGACTTGT 415

Qy 977 AGGTGGGCG 1036
 Db 416 AAGTGGGACACTTGT 475

Qy 1037 ACCCGGCG 1096
 Db 476 ATCCAAATGACTTATCGAGCAAAATGAGCTCCAGAGAGATCCGAGACTTCTCTGTGACGGA 535

Qy 1097 AAGAGTCTGT 1156
 Db 536 AGAGTACTGT 595

Qy 1157 TCACGCTGCGGCG 1216
 Db 596 TCACAGTCTTGTCCAAACGCGAGGTGACACATGTGGGTGAACGAGCAAGCGGTGACATGC 655

Qy 1217 TGCTGTGCGT 1226
 Db 656 TGATGTGTGT 665

RESULT 12

B1404107

LOCUS

B1404107 304 bp mRNA linear EST 14-AUG-2001
 M1-P-CPI-nws-d-12-0-UI-M1-P-CPI Sus scrofa cDNA clone
 M1-P-CPI-nws-d-12-0-UI 3', mRNA sequence.

ACCESSION

B1404107

VERSION

KEYWORDS

SOURCE

ORGANISM

Sus scrofa (p1g)
 Sus scrofa
 Sus scrofa
 Bkaryotes; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
 Mammalia; Euteleostomi; Cetartiodactyla; Suidae; Suidae; Sus.
 1 (bases 1 to 304)
 Bernaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 311)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
PUBMED
8889548
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 39-65,
>GC rich#low complexity.
Seq primer: M13 Forward.
Location/Qualifiers
1. 311
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-G0-up-c-06-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_1lb="UI-R-G0"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; The UI-R-G0
library is a normalized library constructed from a
mixture of rat tissues (nodose ganglia, dorsal root
ganglia, and trigeminal ganglia). The tag is a string of
6 nucleotides present between the Not 1 site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996."

BASE COUNT 31 a 110 c 125 g 45 t

ORIGIN

Query Match 14.5%; Score 242.2; DB 9; Length 311;
Best Local Similarity 86.2%; Pred. No. 2.5e-32;
Matches 269; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

D 943 CGCCCGCTGGCGCCCGCGAGAGCTTCTGTGTGAGGTGGCGCTCCGGGCTGGCGCG 1002
1 CGGCCGTGGCGCCCGCGAGAGCTTCTGTGTGAGGTGGCGCGCCCGCGGCTGGCGCG 60

D 1003 CCGCGCGCTGGCGCTTCCGATCATCTCTGTGCGATCTGTGCGCGCTGCGCGCATTCGAG 120
61 CCGCGAGCTGTGCGCTTCCGATCATCTCTGTGCGATCTGTGCGCGCTGCGCGCATTCGAG 120

D 1063 CTGCGCGCGAGCCGAGCGCGCTGTGATCGCGAGAAAGTACTGGGTGTGGCGCGCGCC 1122
121 CTGCGCGCGATCTCCGTGGCTGTGATCGCGAGAAATCTGGGTGTGGCGCGCGCC 180

D 1123 GAGCGCGGTGGCGAGCGCGCGAGCGCTGATCGCTTCACTGCTGGCGCGCGCGAGCGT 1182
181 GAGCGCGGTGGCGAGCGCGCGAGCGCACTGCTTCACTGCTGCGAGCGCGCGAGCGT 240

D 1183 CTCTGGGAGTCAACGGGCGCTCGCGCGCGCGCTGCTGTGCTGCAACACAGCGAGCG 1242
241 CTGCTGGCGGTGAACGGGCGCGCGCGCGAGCGCTGCTGTGCTGCAACACAGCGAGCG 300

D 1243 CTCTGGGCGCTT 1253
|||||

Db 301 CTCTGGGCGCTT 311

RESULT 15
LOCUS AZ953956
DEFINITION 2M0219K24F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0219K24 F, genomic survey sequence.
ACCESSION AZ953956
VERSION AZ953956.1 GI:13825183
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 362)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, B., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plates: 0219 row: K column: 24
Seq primer: CGTGTAAACGACGCGCCAGT
Clase: plasmid ends
High quality sequence stop: 362.
Location/Qualifiers
1. 362
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0219K24"
/sex="Female"
/lab_host="R. coli strain XL10-Gold, T1-resistant, F-"
/clone_1lb="Mouse 10kb plasmid UUGC2M library"
/note="Vector: pMD42uv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (gi|4732114.9b|AF12072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptor complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent R. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 43 a 136 c 119 g 64 t

ORIGIN

Query Match 13.1%; Score 220.2; DB 28; Length 362;
Best Local Similarity 80.8%; Pred. No. 1.6e-28;
Matches 269; Conservative 0; Mismatches 63; Indels 1; Gaps 1;

D 802 CGGCCGCGCTGTGGCGCGGCGTACTGAGGCGCA-CTGCGCGCTTCAGCGCAACAGCGCG 860

